December 1955

BIOMETRICS

Vol. II No. 4

JOURNAL OF THE BIOMETRIC SOCIETY

Partially Replicated Latin Squares — *W. J. Youden and J. S. Hunter*

The Relative Size of the Inter- and Intra-Block Error in an Incomplete Block Design — *W. A. Thompson, Jr.*

A Note on Design and Analysis of Soil Insecticide Experiments — *P. Sprent*

Covariance Analysis as an Alternative to Stratification in the Control of Gradients — *Anne D. Outhwaite and A. Rutherford*

On the Analysis of Variance of a Two-Way Classification with Unequal Sub-Class Numbers — *Clyde Young Kramer*

A Method of Analysis for a Double Classification Arranged in a Triangular Table — *Nell Ditchburne*

Appropriate Scores in Bio-Assays using Death-Times and Survivor Symptoms — *Johannes Ipsen*

Matrices in Quantal Analysis — *P. J. Claringbold*

The Biometric Society

BIOMETRICS

FOUNDED BY THE BIOMETRICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION

TABLE OF CONTENTS

Partially Replicated Latin Squares
W. J. Youden and J. S. Hunter 399

The Relative Size of the Inter- and Intra-Block Error in an Incomplete Block Design . . . . . W. A. Thompson, Jr. 406

A Note on Design and Analysis of Soil Insecticide Experiments
P. Sprent 427

Covariance Analysis as an Alternative to Stratification in the Control of Gradients, Anne D. Outhwaite and A. Rutherford 431

On the Analysis of Variance of a Two-Way Classification with Unequal Sub-Class Numbers . . . . Clyde Young Kramer 441

A Method of Analysis for a Double Classification Arranged in a Triangular Table . . . . . Nell Ditchburne 453

Appropriate Scores in Bio-Assays using Death-Times and Survivor Symptoms . . . . . Johannes Ipsen 465

Matrices in Quantal Analysis . . . . . P. J. Claringbold 481

Letter to the Editor 502

Queries 504

Change in Editor 506

Abstracts 508

Note and Announcement, News of Members 519

International Biometric Symposium, Campinas, Brazil 521

The Biometric Society 556

Table of Contents and Index to Volume 11 563

Number 4 **December 1955** **Volume 11**

Material for *Biometrics* should be addressed to Dr. J. W. Hopkins, National Research Council, Ottawa 2, Canada, except that authors residing in one of the following organized regions can expedite the handling of their papers by submitting them to the Assistant Editor for that region.

British Region: Dr. D. J. Finney, Dept. of Stat., Univ. of Aberdeen, Aberdeen, Scotland; **Australasian Region:** Dr. E. A. Cornish, University of Adelaide, Adelaide, Australia; **French Region:** Dr. Georges Teissier, Faculte des Sciences de Paris, 1 rue V. Cousin, Paris, France.

Material for Queries should go to Professor G. W. Snedecor, Statistical Laboratory, Iowa State College, Ames, Iowa.

Articles to be considered for publication should be submitted in triplicate.

THE BIOMETRIC SOCIETY

General Officers

President, W. G. Cochran; *Secretary-Treasurer*, C. I. Bliss; *Council*, H. C. Batson, L. L. Cavalli-Sforza, G. M. Cox, Georges Darmois, B. B. Day, D. J. Finney, Sir Ronald Fisher, J. H. Gaddum, M. P. Geppert, P. C. Mahalanobis, Donald Mainland, Leopold Martin, M. Masuyama, P. A. Moran, J. Neyman, J. W. Tukey, Frank Yates, W. J. Youden.

Regional Officers

Eastern North American Region: *Regional President*, D. B. Duncan; *Secretary-Treasurer*, A. M. Dutton. British Region: *Regional President*, R. R. Race; *Secretary*, E. C. Fieller; *Treasurer*, A. R. G. Owen. Western North American Region: *Regional President*, W. J. Dixon; *Secretary-Treasurer*, Elizabeth Vaughan. Australasian Region: *Regional President*, E. J. Williams; *Secretary*, W. B. Hall; *Treasurer*, Miss N. Ditchburne. French Region: *Regional President*, Georges Darmois; *Secretary-Treasurer*, Daniel Schwartz. Belgian Region: *Regional President*, Paul Spehl; *Secretary*, Leopold Martin; *Treasurer*, Claude Panier. Italian Region: *Regional President*, C. Barigozzi; *Secretary*, L. L. Cavalli-Sforza; *Treasurer*, R. Scossiroli. German Region: *Regional President*, E. Ullrich; *Secretary-Treasurer*, W. Ludwig.

National Secretaries

Denmark, N. F. Gjeddebaek; The Netherlands, E. van der Laan; India, V. G. Panse; Japan; M. Hatamura; Switzerland, Arthur Linder; Sweden, H. O. A. Wold; Brazil, Americo Groszmann.

Editorial Board

Biometrics

Editor: Gertrude M. Cox; *Editor-elect:* J. W. Hopkins; *Assistant Editors and Committee Members:* C. I. Bliss, Irwin Bross, E. A. Cornish, W. J. Dixon, Mary Elveback, Ralph Bradley, D. J. Finney, S. Lee Crump, Leopold Martin, K. R. Nair, Horace W. Norton, G. W. Snedecor and Georges Teissier. *Managing Editor:* J. W. Hopkins.

The Biometric Society is an international society devoted to the mathematical and statistical aspects of biology. Biologists, mathematicians, statisticians and others interested in its objectives are invited to become members. Through its regional organizations the Society sponsors regional and local meetings. National secretaries serve the interests of members in Brazil, Denmark, India, Japan, The Netherlands and Sweden, and there are many members at large. Rates (in U.S.A. currency) for full membership in the Society for 1955, including dues for a subscription to BIOMETRICS, are the following: for residents of Canada and the United States $7.00, for members resident in other parts of the world $4.50. Members of the American Statistical Association who are currently subscribing to BIOMETRICS through that organization may become members of The Biometric Society on the payment of $3.00 annual dues if resident in the United States or Canada, and of $1.75 annual dues if resident elsewhere. Information concerning the Society can be obtained from the Secretary-Treasurer, The Biometric Society, Box 1106, New Haven 4, Connecticut, U.S.A.

The annual subscription to BIOMETRICS for non-members of The Biometric Society is $7.00, payable to the Managing Editor, BIOMETRICS, P. O. Box 5457, Raleigh, North Carolina, U.S.A. Members of the American Statistical Association may subscribe to BIOMETRICS through the Secretary of the American Statistical Association at $4.00 per annum.

Second-class mail privileges authorized at New Haven, Conn. Additional entry at Richmond, Va. Business Office, 52 Hillhouse Ave., New Haven, Conn. *Biometrics* is published quarterly—in March, June, September and December.

PARTIALLY REPLICATED LATIN SQUARES

W. J. YOUDEN* AND J. S. HUNTER**
Institute of Statistics, N. C. State College
Raleigh, N. C.

Latin squares have been used for over 30 years in agricultural field trials (1). Often the double restriction imposed by the row and column arrangement brought a welcome reduction in the mean square for error. The row and column mean squares were of no interest. Eventually the Latin square arrangement found application where the rows and columns corresponded to clearly defined physical entities. Youden (3) in studies on tobacco mosaic virus observed that the number of lesions produced on leaves depended upon the position of the leaf on the plant and that plants also differed markedly in susceptibility to lesions. Furthermore, for a given lot of plants, the effect of leaf position was closely the same from one plant to another. Very large reductions in the error mean square resulted from the use of a Latin square. The plants were columns and leaf positions were rows. Here there was some interest in the mean squares for rows and columns though the chief concern was with the treatments applied to the leaves. Later Yates (2) introduced confounding into Latin square designs.

Probably it was inevitable that the Latin square arrangement would be tried when the rows and columns were used for factors that not only were likely to interact with the treatments (letters) but also with each other. The form or appearance of a Latin square remains but the substance is lost. The fractional replication of a factorial experiment that results from this practice is a particularly unfortunate one as all

*On leave from the National Bureau of Standards, Washington 25, D. C.
**Now with the American Cyanamid Company, New York, N. Y.

main effects are directly confounded with two factor interactions and the residual (in 3×3 squares or larger) becomes a snare for the unwary. The plausible appearance of getting something for nothing has trapped and will continue to trap novices in the use of experimental design. For this reason there may be some advantage in making available a slight extension of the Latin square that will give an indication as to whether or not the usual requirement of additive effects for rows, columns and treatments has been met.

The proposed extension consists in performing for a $k \times k$ Latin square k additional experiments so chosen that each row, column and letter enters into $k + 1$ measurements. For convenience the duplicated cells may be shown lying along a principal diagonal. Randomizing the rows and columns will change the pattern of duplicated cells.

	c_1	c_2	c_3
r_1	A	B	CC'
r_2	C	AA'	B
r_3	BB'	C	A

	c_1	c_2	c_3	c_4	c_5	c_6
r_1	AA'	C	F	E	D	B
r_2	D	BB'	E	F	C	A
r_3	B	D	CC'	A	F	E
r_4	E	F	A	DD'	B	C
r_5	F	A	B	C	EE'	D
r_6	C	E	D	B	A	FF'

	c_1	c_2	c_3	c_4
r_1	AA'	C	B	D
r_2	B	D	A	CC'
r_3	D	BB'	C	A
r_4	C	A	DD'	B

It is not difficult to construct various arrangements for each size Latin square. Enough symmetry is retained to lead to reasonably convenient sets of estimates for row, column and treatment effects. The analysis of variance for a $k \times k$ square follows:

TABLE 1

Analysis of variance for partially replicated Latin squares.

Item	d.f.	Sum of squares
Rows	$(k-1)$	$\sum_{i=1}^{k} \frac{R_i^2}{k+1} - \frac{G^2}{k(k+1)}$
Cols. adjusted for rows	$(k-1)$	$\frac{k+1}{k(k+2)} \sum_{i=1}^{k} (C_i')^2$
Treats. adj. for rows, cols.	$(k-1)$	$\frac{k+2}{k(k+3)} \sum_{i=1}^{k} (T_i'')^2$
"Interaction"	$(k-1)(k-2)$	By difference
Error (duplicates)	k	$\frac{1}{2} \sum \left(\text{differences of duplicates}\right)^2$
Total sum of squares	$k(k+1)-1$	$\sum_{i,j=1}^{k} y_{ij}^2 - \frac{G^2}{k(k+1)}$

y_{ij} = an observation in the ith row, jth column
G = Grand Total R_i = Total of ith row
C_i = Total of ith column T_i = Total of ith treatment

$$C_i' = C_i - \frac{R+G}{k+1}$$

where R is that row total associated with column i in which the duplicate occurs.

$$\sum_{i=1}^{k} C_i' = 0$$

$$T_i'' = T_i - \frac{R+C+G}{k+2}$$

where R and C are the row and column totals for row and column in which treatment i is duplicated.

$$\sum_{i=1}^{k} T_i'' = 0$$

Adjusted treatment = mean =

$$t_i = \frac{(k+2)}{k(k+3)} T_i'' + \frac{6}{k(k+1)}$$

Variance difference between treatment means

$$= V(t_i - t_j) = 2\frac{(k+2)}{k(k+1)}s^2$$

As an example consider some data obtained on the density of small bricks. Three different sizes of powder particles were used. Each size was compressed under three pressures and fired at three temperatures. Finally the 27 combinations of this $3 \times 3 \times 3$ factorial were run in duplicate. These data are given in full and then portions of them used to show what happens when a subset of 9 in the form of a Latin square are chosen on the presumption that useful information will be obtained. The same 9 values will then be supplemented by the three duplicates required for the partially replicated Latin square.

The complete set of 54 measurements is given in Table 2. The values in italics (using only the first when both are in italics) are those used in a Latin square selection. The analysis of variance for these 9 results is listed in Table 3 alongside the mean squares for the complete set of data. It is abundantly clear that no useful interpretation is possible using the 9 results nor is there any way to ascertain that the error variance is, in fact, a good deal smaller than the residual mean square of 2258.

TABLE 2

Densities of Briquettes formed from three sizes of particles, compressed at three pressures and fired at three temperatures. The decimal points are omitted. Duplicate values are separated by commas.

Size	Pressure	Temperature, degrees Fahrenheit		
		1900	2000	2300
5–10	5.0	945, *961	*933*, 968	962, *950
	12.5	969, 960	944, 882	*942*, *958*
	20.0	*964*, *964	949, 964	965, *974
10–15	5.0	905, 897	969, 927	*908*, 892
	12.5	*936*, 946	925, 985	892, 904
	20.0	940, 924	*905*, *948*	950, 917
15–20	5.0	*842*, **845*	848, 872	851, *881
	12.5	868, 790	*981*, 989	872, 879
	20.0	845, *880	993, 1020	*890*, *902

TABLE 3

Analysis of Variance for Density Data.

Item	All 54 results		Latin square 9 results	
	d.f.	M.S.	d.f.	M.S.
Temp.	2	6011	2	676
Size	2	17065	2	1404
Pres.	2	3946	2	2598
$T \times S$	4	6834		
$T \times P$	4	466	2	2258
$S \times P$	4	1798		
$T \times S \times P$	8	2170		
Dupli.	27	456		

TABLE 4

Analysis of variance for partially replicated Latin square as explained in Table 1. Data used are the italicized values in Table 2.

Item	d.f	S.S.	M.S.
Size	2	7175.166	
Temp. corr. for size	2	2737.233	
Pres. corr. for size and temp.	2	6662.600	3331
"Interaction"	2	4083.416	2042
Error	3	854.500	285
Total	11	21512.916	

If the experimenter wants to take a chance that the interactions are small the experiment should furnish the means of demonstrating that the gamble has been won. To this end the other three values underlined in Table 2 are combined with the 9 values and the 12 results examined by the formulas in Table 1. Table 4 shows the analysis of variance, the pressures taking the role of treatments. The three additional measurements give a valid estimate of the experimental error provided the order was suitably randomized. The small value for the error variance relative to the interaction mean square gives warning that the interactions are not negligible and that the main effects may not be what they seem to be. It is of course recognized that not all possible interaction effects are measured by this technique, but that these interactions are in fact sampled by the degrees of freedom available.

The sum of squares for temperature when corrected for size and pressure is 2273.4. The sum of squares for size, when corrected for temperature and pressure, is 3554.6. These will be needed provided the duplicates give assurance of little or no interaction. The computation of these corrected sums of squares may be obtained by permuting the roles of R_i, C_i and T_i in the formulas shown in Table 1.

There is another alternative to the one third replicate of the $3 \times 3 \times 3$ factorial if there exists any strong misgivings about the absence of interactions. In Table 2 eight of the entries are marked with an asterisk. These entries constitute a $2 \times 2 \times 2$ experiment using the lowest and highest temperatures, smallest and largest particles, and the lowest and highest pressures. The analysis of variance for these eight results is shown in Table 5.

TABLE 5

Analysis of variance for $2 \times 2 \times 2$ factorial.

Item	d.f.	M.S.
Size	1	20808
Temperature	1	648
Pressure	1	512
$T \times S$	1	162
$T \times P$	1	50
$S \times P$	1	50
$T \times S \times P$	1	338

With all the limitations of single degrees of freedom the experimenter can feel fairly confident that particle size is important. The experiment fails to reveal interactions that appear to arise with the intermediate levels of the factors that were omitted. Altogether the example shows that skimping of the measurements cuts down on the information. Experimenters know that small or even moderate effects are likely to be missed when only a few measurements are made. Furthermore the use of a Latin square arrangement gave results that are likely to be dangerously misleading to workers with little experience in the design and analysis of experiments. For that matter, even the analysis for the complete experiment may mislead the beginner. The small mean square for $T \times P$ may lead to the hasty conclusion that these factors do not interact. But if the three temperature-pressure tables are examined separately for each size of particle, significant interactions will be found. In the analysis of the whole set of data these interactions are compen-

sating. The large mean square for the three factor interaction gives warning of this possibility.

In summary, the experimenter is not always aware that additivity of rows, columns and treatments is a basic assumption for the Latin square. The experimenter sees only that, by identifying rows, columns, and letters with experimental factors, a small subset of treatments is specified. Ultimately the experimenter may learn that there is no unambiguous interpretation of these so called Latin squares unless he has information about the experimental error. The slightly replicated Latin square directs attention to the need for this estimate of error. The degrees of freedom for error are few. On the other hand the duplicates have been chosen to facilitate the examination of the data.

REFERENCES

1. Fisher, R. A., *Statistical Methods for Research Workers.* 1st ed. 1925 Oliver and Boyd. Edinburgh, England.
2. Yates, F., The Design and Analysis of Factorial Experiments. *Techn. Comm. No. 35,* 1935. *Imper. Bur. of Soil Science,* Harpenden, England.
3. Youden, W. J. and H. P. Beale, A statistical study of the local lesion method for estimating tobacco mosaic virus. *Contributions from Boyce Thompson Institute 6,* 437. 1934.

THE RELATIVE SIZE OF THE INTER- AND INTRA-BLOCK ERROR IN AN INCOMPLETE BLOCK DESIGN*

W. A. THOMPSON, JR.
Virginia Polytechnic Institute

1. INTRODUCTION

Results of scientific experiments are frequently classified according to the factors contributing to the observed data. Thus, several technicians may test a property of a certain type of shoe on several different walking courses, and y_{ij} is the score given to the shoe by the ith technician on the jth course. The shoe to shoe factor for a given type shoe is assumed negligible. The two factors considered in this example are then (i) different technicians and (ii) different courses.

An individual technician's gradings will not be completely reproducible, i.e., if he repeatedly grades the same type shoe on the same course he will not always give it the same score. He will tend to give scores which fall in a regular manner about the actual value for the particular shoe and course. This variation is then the error contribution to the score.

If, as is usually the case, we wish to draw inferences about locations other than those where we perform our experiments, it may be more reasonable to assume that the course effect is composed of a large number of effects which have a tendency to counterbalance each other. This suggests to the author that we should also assume the course effects to be random ones, and we will then have two sources of variability with which to deal: (i) the variability in performance due to the different courses and (ii) the variability due to all other factors which we have called the error variability.

*This work was supported in part by the U.S. Army Quartermaster Research and Development Command under Contract No. DA44-109-qm-1488. The views and conclusions in this report are those of the author and do not necessarily reflect the views, or have the indorsement, of the Department of Defense.

Hence, if the ith technician works on the jth course we will assume that

$$y_{ij} = \alpha + \tau_i + \beta_j + \epsilon_{ij}$$

α being an average effect over all technicians and courses, τ_i the effect of the ith technician and β_j the effect of the jth course; ϵ_{ij} is the error contribution. The β_j's are assumed to be independent observations from a normal distribution with mean zero and variance σ_1^2 ; the ϵ_{ij}'s are assumed to be independent and identically distributed with mean zero and variance σ^2.

The reader may recognize that we are dealing with a "mixed" model since the β's are assumed to be random and the τ's fixed.

We now wish to investigate the size of σ_1^2 relative to σ^2. We do this by testing hypotheses concerning σ_1^2/σ^2 $(= \mu)$ or finding a confidence interval for this ratio. It is well known that for a randomized block design this can be done by using the F table. The purpose of this paper is to indicate the further refinements necessary for handling an incomplete block design. The method used here was proposed by Wald (5) and developed by the author (4).

In Section 2, in order to present the reader with a concrete illustration of this method, we shall enlarge upon the shoe example. In Section 3, we shall present, for the use of the experimenter, a general set of operating rules. In Section 4 we provide an illustration of the use of the rules given in Section 3 and discuss the possibility of designing experiments with respect to blocks as well as with respect to treatments.

2. EXAMPLE

The illustration which we shall carry over from 1 is furnished by a military experiment in which we test a type of shoe in order to determine its desirability under various walking conditions.

2.1 *Background and Data for the Illustration.*

In this example a number of technicians are asked to rate a particular kind of shoe according to the following scale:

1. Extremely unsatisfactory
2. Very unsatisfactory
3. Moderately unsatisfactory
4. Slightly unsatisfactory
5. Not good, not bad
6. Slightly satisfactory

7. Moderately satisfactory
8. Very satisfactory
9. Extremely satisfactory

Thus, the technicians will give the higher grades to the shoes which they feel are more desirable.

Let us suppose that in this experiment there are six different courses on which the technicians will judge the shoes; these courses are assumed to be randomly chosen from the locations at which it is likely that the shoes will be used.

If we use twenty technicians and six courses, we could let each technician go over each course. We would then have 120 different ratings or y_{ij}'s. However, it may be too expensive or inconvenient to get as many as 120 different ratings, and 60 may be the largest number of ratings that is feasible. Of course, the more ratings that are made, the more accurate will be the knowledge gained from the experiment. On the other hand, there are other ways besides enlarging the number of ratings which will increase the information obtained; particularly, the design of the experiment will insure that resources are used to the best advantage.

In the case at hand a very worthwhile design requiring 60 different ratings would be the following one

course \ technicians	1 5 9 13 17	2 6 10 14 18	3 7 11 15 19	4 8 12 16 20
1	X	X		
2	X		X	
3	X			X
4			X	X
5		X		X
6		X	X	

That is, technicians 1, 5, 9, 13 and 17 go over courses 1, 2, 3; technicians 2, 6, 10, 14 and 18 go over courses 1, 5 and 6, etc. [This is design SR6 Bose, Clatworthy, and Shrikhande (2).]

Using the method of scoring and the design indicated above, the following table contains the raw scores of the actual experiment.

course / technician	1	2	3	4	5	6	Total	Mean
1	3	8	9				20	6.67
2	1				1	5	7	2.33
3		9		4		4	17	5.67
4			9	5	3		17	5.67
5	3	6	6				15	5
6	4				7	1	12	4
7		2		3		9	14	4.67
8			5	6	9		20	6.67
9	6	2	9				17	5.67
10	4				5	9	18	6
11		6		8		1	15	5
12			9	5	8		22	7.33
13	6	5	1				12	4
14	2				6	6	14	4.67
15		4		8		7	19	6.33
16			9	5	2		16	5.33
17	5	9	7				21	7
18	5				1	3	9	3
19		9		6		7	22	7.33
20			6	7	9		22	7.33
Total	39	60	70	57	51	52	329	

Thus, technician 1 rated the shoe to be moderately unsatisfactory on course 1, very satisfactory on course 2 and extremely satisfactory on course 3.

A property of this design which is worth noting is that a number of the technicians do exactly the same things. This makes it possible to instruct and to transport the men in groups and so carry out the experiment with a minimum of confusion. The advantage of transporting the technicians in groups becomes apparent when one alters the circumstances of the present problem and considers the possibility of having 4 groups of 5 observers each watch six different maneuvers.

We have indicated that we are interested in determining the size of

$$\mu = \sigma_1^2/\sigma^2 = \frac{\text{variation due to courses}}{\text{variation due to error}}.$$

In order to show why we might be interested in this information let us outline some of the possible actions which we might take as a result of knowing σ_1^2/σ^2 exactly.

If the course variability is relatively small, then some of the possible

actions which might result as a consequence of having run the experiment are:

1) In analysing the results of the experiment we will "pool" the data over all courses to determine how satisfactory the shoe being tested is. That is, we treat the data as though it was taken all on the same course instead of on six different ones.

2) If the shoe being tested has an overall grade of satisfactory, then we may recommend that the army use this shoe in all locations represented in the test. Hence we "standardize" on this particular shoe.

If, however, we find σ_1^2/σ^2 large, then we will be unable to standardize on a particular shoe since our experiment will have indicated that the desirability of the shoe is not constant over the locations of expected use. A re-examination of the testing procedure and investigation of the properties of the shoe may be indicated.

2.2 *Some Reasons.*

The reader may have a better insight for what is going on if we sketch very briefly some of the reasons for analysing the experiment as we shall in the next section.

In order to remove the effect of the different technicians we average them out by taking the adjusted course totals.

2.2.1 p_i = adjusted course total for course i
= ith course total—sum of means of technicians who are judging on the ith course.

In some designs and for fixed β's the p's would be estimates of the course effects, but in the present one for given β's

2.2.2
$$E(p_1) = \frac{20}{3}\beta_1 - \frac{5}{3}(\beta_2 + \beta_3 + \beta_5 + \beta_6)$$

$$E(p_2) = \frac{20}{3}\beta_2 - \frac{5}{3}(\beta_1 + \beta_3 + \beta_4 + \beta_6)$$

$$E(p_3) = \frac{20}{3}\beta_3 - \frac{5}{3}(\beta_1 + \beta_4 + \beta_2 + \beta_5)$$

$$E(p_4) = \frac{20}{3}\beta_4 - \frac{5}{3}(\beta_2 + \beta_5 + \beta_3 + \beta_6)$$

$$E(p_5) = \frac{20}{3}\beta_5 - \frac{5}{3}(\beta_1 + \beta_4 + \beta_3 + \beta_6)$$

$$E(p_6) = \frac{20}{3}\beta_6 - \frac{5}{3}(\beta_1 + \beta_2 + \beta_4 + \beta_5)$$

or $E(p) = D\beta$ where β_i is the ith course effect and D is the matrix of coefficients of the β's.

The unconditional variance-covariance matrix of $p_1 , p_2 , \cdots , p_6$ is $D\sigma^2 + D^2\sigma_1^2$.

At this point we choose a set of orthogonal linear functions of $p_1 , \cdots , p_6$, say $z_1 , z_2 , \cdots , z_6$. $z_1 , z_2 , \cdots , z_6$ may be chosen so that $z_1 , z_2 , \cdots , z_5$ have the distribution:

$$\text{constant exp}\left(-\frac{1}{2}\sum_{i=1}^{5}\frac{z_i^2}{e_i\sigma^2 + e_i^2\sigma_1^2}\right) \qquad dz_1 \cdots dz_5$$

while z_6 is zero with probability one. For the mathematical reader $e_1 , \cdots , e_5$ are the non-zero characteristic roots of D.

We now consider the ratio

$$\left(\frac{1}{5}\sum_{i=1}^{5}\frac{z_i^2}{e_i\sigma^2 + e_i^2\sigma_1^2}\right)\Big/\frac{1}{\sigma^2} \qquad \text{(M.S.E.)}$$

where, of course, M.S.E. stands for mean square due to error.

We may prove that the numerator and denominator of this ratio are independently distributed according to the Chi-Square distribution, and hence the ratio has the F distribution. We denote this ratio by $F(\mu)$ where $\mu = \sigma_1^2/\sigma^2$.

2.2.3 $$F(\mu) = \frac{1}{5(\text{M.S.E.})}\sum_{i=1}^{5}\frac{z_i^2}{e_i + e_i^2\mu} .$$

Wald (5) points out that $F(\mu)$ is a decreasing function of μ, and hence, if

$$\Pr[F(\mu) < F_{1-\alpha}] = 1 - \alpha,$$

then

2.2.4 $$\Pr(\mu > \mu_{1-\alpha}) = 1 - \alpha,$$

where $\mu_{1-\alpha}$ is defined by $F(\mu_{1-\alpha}) = F_{1-\alpha}$ and where $F_{1-\alpha}$ is the ordinate of the cumulative distribution function of an F statistic with the appropriate number of degrees of freedom. We will then have placed an upper confidence bound on μ with a probability of $1 - \alpha$. We may also find a lower confidence bound for μ and hence a confidence interval for μ (though not a unique one).

Thus, in order to make confidence statements about $\mu = \sigma_1^2/\sigma^2$, we must solve equations of the type

$$\sum_{i=1}^{5}\frac{z_i^2}{e_i + e_i^2\mu} = a,$$

where $a = F_{1-\alpha}$ (M.S.E.)5. For the present incomplete block design it is true that there are only two different e's (i.e., the matrix D has only two different characteristic roots). We thus have a much easier equation to solve:

2.2.5
$$\frac{\sum_1}{e_1 + e_1^2\mu} + \frac{\sum_2}{e_2 + e_2^2\mu} = a,$$

where

$$\sum_1 = \sum_{\substack{i \\ e_i = e_1}} z_i^2 = \frac{e_1}{e_1 - e_2}\left(\sum_{i=1}^{6} p_i^2 - e_2 \sum_{i=1}^{6} m_i p_i \right)$$

and

$$\sum_2 = \sum_{\substack{i \\ e_i = e_2}} z_i^2 = \frac{e_2}{e_2 - e_1}\left(\sum_{i=1}^{6} p_i^2 - e_1 \sum_{i=1}^{6} m_i p_i \right).$$

Here m_1 , m_2 , $\cdots$, m_6 are found by solving equations 2.2.2 for the β's after having substituted p_i for $E(p_i)$. The evaluation of $\sum_1$ and $\sum_2$ in terms of $p_1 \cdots p_6$ and $m_1 \cdots m_6$ is not immediately evident, but will not be discussed here.

We may now solve equation 2.2.5 by clearing of fractions, evaluating $\sum_1$ and $\sum_2$ in terms of p_1 , $\cdots$, p_6 and considering the solutions of the resulting quadratic equations. The non-extraneous solution to the quadratic is then also a solution to our original equation; it turns out to be

2.2.6
$$\mu_0 = \frac{-d + \sqrt{d^2 - 4bc}}{2b}$$

where

2.2.7
$$\begin{aligned} \Delta^* &= e_1 e_2 \\ H^* &= (e_1 + e_2) \\ a &= F_{1-\alpha}(\text{M.S.E.})5 \\ b &= a e_1 e_2 = a\Delta^* \\ c &= a - \sum_{i=1}^{6} m_i p_i \\ d &= c(e_1 + e_2) + \sum_{i=1}^{6} p_i^2 \\ &= cH^* + \sum p_i^2 \end{aligned}$$

2.3 *The Analysis.*

We will use the shoe data that we have been discussing to illustrate two methods of studying the ratio

$$\mu = \frac{\text{variation due to courses}}{\text{variation due to error}}.$$

From 2.2.1 we compute:

$$p_1 = -9.34, \quad p_4 = -4.33$$
$$p_2 = 2.67, \quad p_5 = -1.33$$
$$p_3 = 9.33, \quad p_6 = 3.00.$$

For the present experiment estimates of the β's if they were fixed would be

$$m_1 = \frac{1}{40}(5p_1 - p_4) = -1.059$$

$$m_2 = \frac{1}{40}(5p_2 - p_5) = .367$$

$$m_3 = \frac{1}{40}(5p_3 - p_6) = 1.091$$

2.3.1

$$m_4 = \frac{1}{40}(5p_4 - p_1) = -.308$$

$$m_5 = \frac{1}{40}(5p_5 - p_2) = -.233$$

$$m_6 = \frac{1}{40}(5p_6 - p_3) = .142$$

It should be stated that checks on the calculations of the p_i's and m_i's are possible by $\sum p_i = 0$ and $\sum m_i = 0$.

Note also that equations 2.3.1 satisfy 2.2.2 where $E(p_i)$ is replaced by p_i.

In the two methods of studying the ratio μ we will be using information which can be systematically calculated in two tables; these two tables represent a splitting up of the total sum of squares in two different ways. Computing instructions for the table entries are included in the table.

One method of studying the ratio μ is as follows:

We may decide that if the variation due to locations is less than 1/10, say, of the total variation on some particular kind of equipment

(a shoe) we won't study that item. Thus we must decide: Is

$$\mu = \frac{\text{variation due to courses}}{\text{variation due to other causes}}$$

greater than 1/10 or less than 1/10? Or, in other words, we test the hypothesis H_0: $\mu \leq 1/10$ vs. $H_1 : \mu > 1/10$.

TABLE 2.3.1
Intra-Block Analysis

Source	Degrees of Freedom	Sum of Squares
Course (adj.)	5	$\sum_{i=1}^{6} m_i p_i = 23.1223$
Man (unadj.)	19	$\frac{\sum_{i=1}^{20} (\text{man total})^2}{3} - \frac{(\text{Grand total})^2}{60} = 116.3100$
Error	35	(subtract course and man S.S. from total) $= 261.5477$
Total	59	$\sum_{i,j} y_{ij}^2 - \frac{(\text{Grand total})^2}{60} = 400.9800$

TABLE 2.3.2
Inter-Block Analysis

Source	Degrees of Freedom	Sum of Squares
Course (unadj.)	5	$\frac{\sum_{i=1}^{6} (\text{course total})^2}{10} - \frac{(\text{Grand mean})^2}{60} = 53.4800$
Man (adj.)	19	subtract course and error S.S. from total $= 85.9553$
Error	35	(transfer from table I) $= 261.5477$
Total	59	(transfer from table I)

We compute that, under the assumption that μ equals 1/10, the numerator of our F statistic (the left hand side of equation 2.2.5 divided by 5) is

$$2.3.2 \qquad \left(\frac{\sum_1}{10 + 10^2\left(\frac{1}{10}\right)} + \frac{\sum_2}{\left(\frac{20}{3}\right) + \left(\frac{20}{3}\right)^2\left(\frac{1}{10}\right)} \right) \frac{1}{5} = 2.4340$$

where 10 ($= e_1$) and 20/3 ($= e_2$) are the characteristic roots of which we spoke in Section 2 and

$$\sum_1 = \frac{10}{10 - \frac{20}{3}} \left\{ \sum_{i=1}^{6} p_i^2 - \frac{20}{3} \begin{pmatrix} \text{adj. sum of} \\ \text{squares due} \\ \text{to courses} \end{pmatrix} \right\}$$

$$= 170.349$$

$$\sum_2 = \frac{\frac{20}{3}}{\frac{20}{3} - 10} \left\{ \sum_{i=1}^{6} p_i^2 - 10 \begin{pmatrix} \text{adj. sum of} \\ \text{squares due} \\ \text{to courses} \end{pmatrix} \right\}$$

$$= 40.6$$

The mean square due to error is 7.47. The ratio of 2.3.2 to the mean square due to error is 2.434/7.47 = .326, which is a very insignificant value for an F variate with 5 and 35 d.f. Hence, the probability of F being this large is quite high under the null hypothesis and we conclude that $\mu \leq 1/10$. We would, therefore, not study the effect of location on shoe satisfaction.

A second method of studying this ratio is to attempt to place a confidence interval on it.

We may state that in our example

$$\frac{\text{variation due to course}}{\text{variation due to error}}$$

is between 0 and .3184 with a probability of .95. The computation of this confidence interval is based on 2.2.4.

We note that

$$\Pr(\mu > \mu_{.975}) = .975$$

$$\Pr(\mu > \mu_{.025}) = .025$$

$$\Pr(\mu_{.975} < \mu < \mu_{.025}) = .975 - .025 = .95.$$

The computation of $\mu_{.025}$ according to the computing procedure of 2.2.7 is:

$$F_{.025} = .161$$

$$a = (.161)(261.5477)(.143) = 6.022$$

$$b = 401.48674$$

$$\sum m_i p_i = 23.1223; \qquad \sum p_i^2 = 210.9312$$

$$d = -285.062 + 210.9312 = -74.1308$$

$$c = -17.1003$$

$$\mu_{.025} = \frac{+74.1308 + \sqrt{32{,}957.5475}}{802.9734} = \frac{74.1308 + 181.5421}{802.9734} = .3184$$

We compute $\mu_{.975}$ in an analogous manner:

$$F_{.975} = 2.93$$

$$a = (2.93)(261.5477)(.143) = 109.586$$

$$b = 7306.09862$$

$$\sum p_i^2 = 210.9312; \qquad \sum m_i p_i = 23.1223$$

$$d = 1441.350 + 210.9312 = 1652.2812$$

$$c = 86.4637$$

$$\mu_{.975} = \frac{-1652.2812 + \sqrt{203{,}183.8939}}{14{,}612.1972} = \frac{-1652.2812 + 450.7593}{14{,}612.1972}$$

$$= -.0822$$

Then $P[\mu_{.975} < \mu < \mu_{.025}] = .95$; however, we know that μ is a ratio of squares and hence is positive. We may thus substitute 0 for μ_2 and find

$$P[0 < \mu < .318] \geq .95.$$

3. THE GENERAL CASE

3.1 *The Incomplete Block Variance Components Model.*

We now leave our illustration and generalize the computing rules so that the method may be applied to a general class of designs.

We again state our assumptions. We consider y_{ij} ($i = 1, \cdots, v$; $j = 1, \cdots, b$) to be the "yield" from the ith "treatment" and jth "block" of a statistical experiment using an incomplete block design. The reason for the quotes above is to remind the reader that these

terms may refer to applications which are not at all agricultural in nature. We further assume that the y_{ij}'s are independent and normally distributed random variables for given block effects $\beta_1, \cdots, \beta_b$ and that if the ith treatment appears in the jth block, then

$$y_{ij} = \alpha + \tau_i + \beta_j + \epsilon_{ij}$$

and the variance of y_{ij} is σ^2. In addition we assume the β's are independent and identically normally distributed with mean 0 and variance σ_1^2. Note that if the β's were unknown parameters instead of random variables that we would have the general incomplete block model with fixed effects which appears in analysis of variance (see for example Bose (1)). The total number of observations will be denoted by N, the number of treatments in each block by k, and the number of times a treatment is replicated by r. Only designs for which k and r are the same for all blocks and treatments respectively will be considered. We will denote the jth block total by B_j and the ith treatment total by T_i. Then in order to average out the treatment effects we consider the adjusted block totals

$$p_j = \begin{bmatrix} j\text{th} \\ \text{block} \\ \text{total} \end{bmatrix} - \frac{1}{k}\begin{pmatrix} \text{sum of all treatment} \\ \text{totals for treatments} \\ \text{occurring in } j\text{th block} \end{pmatrix} : j = 1, \cdots, b$$

The expectations of the p_j's are then 0 since the block effects are assumed to be random.

3.2 *Linked Block Designs.*

An incomplete block design has been defined to be a linked block design if

3.2.1
(i) Each block has the same number of treatments k,
(ii) Each treatment occurs in r blocks,
(iii) Any two blocks have the same number of treatments λ^* in common.

These designs were used by Youden (6), and are duals of the well known balanced incomplete block designs.

We define

3.2.3
$$e = [k(r - 1) - \lambda^*]/r$$

The illustration of Section 1 is not a linked block design but is partially linked (to be discussed in 3.3); nevertheless in a manner analogous to the illustration of 2, we may show that if F_α is the value of an F variate which has ordinate α and degrees of freedom $b - 1$ and $N - v - b + 1$,

then

$$3.2.4 \qquad \mu_\alpha = \frac{1}{e^2 F_\alpha} \cdot \frac{N - v - b + 1}{b - 1} \cdot \frac{\sum p_i^2}{\text{S.S.E.}} - \frac{1}{e}$$

is a lower bound for μ with probability α. S.S.E. is, of course, the sum of squares for error.

We may systematize the computation of $\sum p_i^2$ and S.S.E. in the following table:

TABLE 3.2.1

Source of Variation	d.f.	S.S.
Blocks eliminating treatments	$b - 1$	$\frac{r}{b\lambda^*} \sum p_i^2$
Treatments ignoring blocks	$v - 1$	$1/r \sum T_i^2 - (\sum y_i)^2/N$
Error	$N - b - v + 1$	S.S.E. (by subtraction)
Total	$N - 1$	$\sum_i y_i^2 - (\sum y_i)^2/N$

The reader may recognize this as being similar to the analysis of variance table for balanced incomplete blocks. T_i is the total for the ith treatment. We may now state the following:

Rules for linked block designs.

step i. Compute e from 3.2.3

step ii. *Confidence Intervals*

a. compute μ_{α_1} and μ_{α_2} from 3.2.4 and table 3.2.1

b. If $\alpha_1 < \alpha_2$, then $\mu_{\alpha_2} < \mu < \mu_{\alpha_1}$ is a confidence interval for μ with confidence coefficient $\alpha_2 - \alpha_1$. $\mu_{\alpha_2} < \mu$ is a confidence region for μ with confidence coefficient α_2; and $\mu < \mu_{\alpha_1}$ is a confidence region for μ with confidence coefficient $1 - \alpha_1$.

step ii′. *Size α test of $\mu \leq \mu_0$ vs $\mu > \mu_0$*

accept $\mu \leq \mu_0$ if $F(\mu_0) = \frac{N - b - v + 1}{b - 1} \cdot \frac{1}{e + e^2\mu_0} \cdot \frac{\sum p_i^2}{\text{S.S.E.}}$

is less than F_α and accept $\mu > \mu_0$ otherwise.

3.3 *Partially Linked Designs.*

The illustration of Section 2 is a partially linked design and we here generalize the results of that section.

The dual of a Partially Balanced Incomplete Block Design is obtained by interchanging the roles of the treatments and blocks. In analogy with linked block designs we may call these dual designs partially linked designs. The following conditions are satisfied:

(i) The experimental material is divided into b blocks of k units each, different treatments being applied to the units in the same block.

(ii) There are v treatments, each of which occurs in r blocks.

(iii) There can be established a relation of association between any two blocks satisfying the following requirements:

a) Two blocks are either 1st, 2nd, $\cdots$, or m^*th associates.

b) Each block has exactly n_i^*, ith associates ($i = 1, 2, \cdots, m^*$).

c) Given any two blocks which are ith associates, the number of blocks common to the jth associates of the first, and the kth associates of the second is p_{jk}^{i*} and is independent of the pair of blocks with which we start. Also $p_{jk}^{i*} = p_{kj}^{i*}$.

(iv) Two blocks which are ith associates contain exactly λ_i^* common treatments.

If the dual of the design we are working with is tabulated in Bose, Clatworthy and Shrikhande (2) then H^* and Δ^* are the H and Δ tabulated there for this dual design. In the present framework, the key result for partially linked designs is that if F_α is the value of an F variate which has ordinate α and degrees of freedom $b - 1$ and $N - v - b + 1$; and if

$$a_\alpha = F_\alpha \begin{bmatrix} \text{sum of} \\ \text{squares} \\ \text{due to} \\ \text{error} \end{bmatrix} \frac{b - 1}{N - b - v + 1}$$

$$b_\alpha = a_\alpha \Delta^*$$

3.3.1

$$c_\alpha = a_\alpha - \begin{bmatrix} \text{adjusted} \\ \text{sum of} \\ \text{squares for} \\ \text{blocks} \end{bmatrix}$$

$$d_\alpha = c_\alpha H^* + \sum p_i^z$$

then

$$3.3.2 \qquad \mu_\alpha = \frac{-d_\alpha + \sqrt{d_\alpha^2 - 4b_\alpha c_\alpha}}{2b_\alpha}$$

is a lower bound for μ with probability α.

We systematize the computation of the adjusted sum of squares for blocks, $\sum p_i^2$, and S.S.E. in the following two tables.

TABLE 3.3.1

Source	d.f.	Sum of Squares
Blocks (adj.)	$b-1$	$\sum_{i=1}^{b} m_i p_i$
Treatments (unadj.)	$v-1$	$\sum_{i=1}^{v} \frac{T_i^2}{r} - \frac{G^2}{N}$
Error	$N-b-v+1$	S_e^2 (by subtraction)
Total	$N-1$	$\sum_{i,j} y_{ij}^2 - \frac{G^2}{N}$

TABLE 3.3.2

Source	d.f.	Sum of Squares
Blocks (unadj.)	$b-1$	$\frac{1}{k}\sum_{j=1}^{b} B_j^2 - \frac{G^2}{N}$
Treatments (adj.)	$v-1$	S_t^2 (by subtraction)
Error	$N-v-b+1$	S_e^2 (by transfer from table 3.3.1)
Total	$N-1$	$\sum y_{ij}^2 - \frac{G^2}{N}$

Here $m_1, m_2, \cdots, m_b$ are a solution to the equations

$$\begin{aligned} d_{11}m_1 + d_{12}m_2 + \cdots + d_{1b}m_b &= p_1 \\ d_{12}m_1 + d_{22}m_2 + \cdots + d_{2b}m_b &= p_2 \\ \cdot \quad & \quad \cdot \quad \quad \cdot \quad \cdot \\ \cdot \quad & \quad \cdot \quad \quad \cdot \quad \cdot \\ \cdot \quad & \quad \cdot \quad \quad \cdot \quad \cdot \\ d_{1b}m_1 + d_{2b}m_2 + \cdots + d_{bb}m_b &= p_b \end{aligned}$$

Hence $m_1 - m_2$ would estimate $\beta_1 - \beta_2$ if the β's were fixed instead of random. G is the grand total of the y_{ij}'s.

We summarize:

rules for partially linked designs ($m^* = 2$)

i. *Two sided confidence interval*
   a. If possible, look up the dual of the design you are using in the B.C.S. catalog (2). Set H and Δ found there equal to H^* and Δ^* respectively.
   b. Compute S.S.E. and $\sum m_i p_i$ from tables 3.3.1 and 3.3.2.
   c. Compute a_{α_1}, b_{α_1}, d_{α_1}, c_{α_1} and μ_{α_1} from 3.3.1 and 3.3.2 in turn. Do the same for μ_{α_2}.
   d. If $\alpha_1 < \alpha_2$, then $\mu_{\alpha_2} < \mu < \mu_{\alpha_1}$ is a confidence interval for μ with confidence coefficient $\alpha_2 - \alpha_1$.

ii. *One sided confidence interval*
   a.
   b. } same as in *i*.
   c.
   d. $\mu_{\alpha_2} < \mu$ is a confidence region for μ with confidence coefficient α_2; and $\mu < \mu_{\alpha_1}$ is a confidence region for μ with confidence coefficient $1 - \alpha_1$.

iii. *Test of* $\mu \leq \mu_0$ *vs* $\mu > \mu_0$
   accept $\mu \leq \mu_0$ if

$$\frac{N - b - v + 1}{b - 1} \cdot \frac{1}{\text{S.S.E.}} \cdot \frac{(\sum m_i p_i)[1 + \mu_0 H^*] - \mu_0 \sum p_i^2}{1 + \mu_0 H^* + \mu_0^2 \Delta^*}$$

is less than F_α and accept $\mu > \mu_0$ otherwise.

3.4 *A note of warning.*

It is to be emphasized that the above rules apply only to linked block and partially linked designs ($m^* = 2$) whose duals are listed in

the B.C.S. catalog. This latter group, however, includes essentially all known partially linked designs with $m^* = 2$ except the duals of lattices. Our procedure will then be as follows: In setting up the experiment pick a design which is (i) balanced or partially balanced ($m = 2$) and (ii) linked or partially linked ($m^* = 2$); then, if the experiment is carried out intact we may use the rules of this section to analyze the variance of the block effects and the well known procedures of inter-block analysis to analyze the treatment effects.

If we were to use an arbitrary design we would, in general, find it difficult to identify the nature of the dual and hence would not be able to apply the rules of this section.

4. SECOND EXAMPLE

4.1 *Introduction*

It has long been recognized that designing experiments with a view to analyzing the "treatment" effects simplifies the analysis greatly; it isn't surprising then, that if we want to analyze the "block" variability that such an analysis is greatly simplified by designing the experiment with this in mind.

Now, partially balancing a design is a method of designing an experiment to simplify the treatment analysis and it appears from the preceding sections that partially linking is a method of simplifying the block error analysis. Fortunately designs may be both partially balanced and partially linked. One such "nice" design from this point of view will be analyzed to illustrate these general remarks.

4.2 *A Linked Block Design*

As an illustrative example we will give the analysis of an experiment on the yields of 35 varieties of oats by Dr. G. K. Middleton. This is Design 2, Table III A of Bose and Shimamoto (3). The treatment analysis of this experiment may be found in Bose, Clatworthy, and Shrikhande (2). This is a particularly interesting design from the present point of view, because it is both partially balanced and a linked block design. The dual of the Bose and Shimamoto Design 2, Table III A is the balanced incomplete block design with parameters $u = 15$, $b = 35$, $r = 7$, $k = 3$ and $\lambda = 1$, hence we follow the rules of 3.2 using:

$$b = 15 \qquad u = 35$$

$$k = 7 \qquad r = 3$$

$$\lambda^* = 1$$

TABLE 4.2.1

Treat. \ Blocks	1	2	3	4	5	6	7	8
1	442							
2		442					556	
3				407				
4	502		498				580	
5		456				504		
6						515		406
7				526				
8			452		443			
9				440		526		
10			320					
11		482						
12					280			365
13							434	
14					417			
15	314							260
16							491	
17	428							
18	512	427		413				
19		436	378					
20						314	366	
21			272	306				315
22			413			286		
23					265	271		
24								
25				212				
26		431			326			
27								
28							284	285
29								365
30	404				328			
31		411						380
32								
33				385	286		370	
34			334					
35	389					455		
Block Total $= B_i$	2991	3085	2667	2689	2345	2871	3081	2376
$B_i - p_i$	2920	3077	2645	2830	2568	2910	2909	2468
p_i	71	8	22	−141	−223	−39	172	−92

TABLE 4.2.1 (*Cont.*)

9	10	11	12	13	14	15	Treat. Total	Treat. Means	Treat.
	400					438	1280	427	1
		347					1345	488	2
	452			498			1357	452	3
							1580	527	4
				576			1536	512	5
		592					1513	504	6
			642			522	1690	563	7
		496					1391	464	8
					433		1399	466	9
325						345	990	330	10
480			532				1494	498	11
	364						1009	336	12
			376		431		1241	414	13
328				403			1148	383	14
					244		818	273	15
				634		488	1613	538	16
		431		486			1345	448	17
							1352	451	18
	345						1159	386	19
	430						1110	370	20
							893	298	21
			386				1085	362	22
						260	796	265	23
		275			276	312	863	288	24
292		254					758	253	25
					472		1229	410	26
325	322				352		999	333	27
226							795	265	28
			476	418			1259	420	29
			356				1088	363	30
						326	1117	372	31
	425	510	407				1342	447	32
							1041	347	33
				265	235		834	278	34
450							1294	431	35
2426	2738	2905	3175	3280	2443	2691	41,763	13,922	
2493	2751	2852	3067	3031	2462	2783	41,766		
−67	−13	53	108	249	−19	−92	−3		

4.2.1 is a table in which the yields of oats are cross classified according to what treatment they represent and what block they were planted in. The column sums are then the block totals, B_i , and the row sums are the treatment totals, T_i . A check is furnished by summing the block totals and the treatment totals, since both sums should equal the grand total, G. An additional column to the right may be used to calculate the treatment means T_i/k, and this may again be checked by summing and comparing with the grand mean since $\sum T_i/k = G/k$. We may also use this table to compute the p_i's. We first compute $B_i - p_i$ by summing the means of the treatments appearing in the ith column. As an example

$$B_1 - p_1 = 427 + 527 + 273 + 448 + 451 + 363 + 431$$
$$= 2920$$

p_i is then calculated by subtracting $B_i - p_i$ from B_i . $B_1 = 2991 - 2920 = 71$. Checks are furnished at each of the last two stages by resorting to the identities $\sum_i (B_i - p_i) = G$ and $\sum_i p_i = 0$, respectively.

In the present numerical example Table 3.2.1 becomes

TABLE 4.2.2

Source of Variation	d.f.	S.S.
Blocks eliminating treatments	14	$\frac{3}{15.1}$ 205,165
Treatments ignoring blocks	34	753,254
Error	56	141,155
Total	104	935,442

From 3.2.3 we see that

$$e = \frac{1}{r}[k(r-1) - \lambda^*]$$
$$= \frac{1}{3}[(7)(2) - 1]$$
$$= \frac{13}{3}.$$

Let $\alpha_1 = .05$ and $\alpha_2 = .95$, then $F_{\alpha_2} = 1.75$ and $F_{\alpha_1} = 1/2.11 = .475$. Now from 3.2.4 we see that

$$\mu_s = \frac{1}{\left(\frac{13}{3}\right)^2 F_s} \frac{104}{14} \cdot \frac{205{,}165}{141{,}155} - \frac{1}{\frac{13}{3}}; \qquad s = \alpha_1, \alpha_2$$

$$\mu_{\alpha_1} = .097$$

$$\mu_{\alpha_2} = 1.884.$$

And finally (.097, 1.884) is a 90% confidence interval for μ, (0, 1.884) and (.097, ∞) are 95% confidence regions for μ.

If we wish to perform the test of step ii' with $\mu_0 = 3$, say, and significance level .05; then we compute $F(3)$

$$F(3) = \frac{104}{14} \frac{1}{\frac{13}{3} + \left(\frac{13}{3}\right)^2 3} \frac{205{,}165}{141{,}155}$$

$$= .178.$$

Now since $F(3) = .178 < F_{.05} = .475$, we accept the hypothesis that $\mu \leq 3$. It is easily seen that if μ_0 has any value greater than 1.884 then we will accept the hypothesis $\mu \leq \mu_0$.

REFERENCES

1. R. C. Bose, "Least Square Aspects of Analysis of Variance," Institute of Statistics, University of North Carolina.
2. R. C. Bose, W. H. Clatworthy and S. S. Shrikhande, "Tables of Partially Balanced Designs with Two Associative Classes," Institute of Statistics, University of North Carolina.
3. R. C. Bose and T. Shimamoto, "Classification and Analysis of Partially Balanced Incomplete Block Designs with Two Associate Classes," *Journal of the American Statistical Association*, Vol. 47, No. 258.
4. W. A. Thompson, Jr., "Statistical Analysis of Military Risks Resulting from Extreme Environmental Conditions," Report No. 3, 1954.
5. Abraham Wald, "A Note on Regression Analysis," *Annals of Mathematical Statistics*, Vol. 18, p. 586.
6. W. J. Youden, "Linked Blocks: A New Class of Incomplete Block Designs," *Biometrics*, 7 (1951), 124 (abstract).

A NOTE ON DESIGN AND ANALYSIS OF SOIL INSECTICIDE EXPERIMENTS

P. SPRENT
University of Tasmania

In a paper in this journal, van der Reyden [2] proposes a design in which control plots, one adjacent to each treated plot, are used to obtain adjustments for uneven distribution of insects in the soil. He discusses the application of his method to a soil insecticide experiment in rectangular lattice design. It unfortunately proved impossible to obtain the original data from van der Reyden's experiment, so that the present writer was unable to clear up certain points of ambiguity in the necessarily brief description, or to try different alternative methods of analysis. In this note attention is drawn to certain aspects of the method which do not seem satisfactory.

Van der Reyden's method of analysis appears to consist of analysing separately the results for the treated and control plots, obtaining the estimated treatment effects and residuals for each plot. Two corrections are then applied to the observed value for each treatment plot, viz:

(i) the value of the "treatment" constant of the corresponding control plot is subtracted from the treatment constant for the the treated plot,

(ii) the residual of the treated plot is replaced by the mean of the residuals of the treated plot and the corresponding control plot.

"Corrected" values for each treated plot are then reconstructed from the mean, replicate, block and treatment constants and residuals, modified as above. The results are subjected to an ordinary analysis of variance.

With regard to correction (i) there is some ambiguity in van der Reyden's paper. It is implied on p. 293 that the "treatment" constant for the control plot is added rather than subtracted. There is further evidence on this point, to be discussed later, in the numerical example. It suffices to point out here that if we regard these control "treatment" effects as estimates of the infestation on the corresponding treated plots, then what we measure on the treated plots as treatment is a "treatment + infestation" effect. The obvious correction is to subtract our estimate of infestation.

It should be noted that correction (i) corresponds to the correction used in the ordinary analysis of covariance when the regression coefficient, b, is taken as unity. In this case the residuals for the corrected values are $\epsilon_u - \epsilon_w$, and the corresponding error variance is

$$V(\epsilon_u) + V(\epsilon_w) - 2 \text{ cov } (\epsilon_u , \epsilon_w) \tag{1}$$

where ϵ_u and ϵ_w are the residuals for the treated and control plot respectively.

It should be noted that the appropriate residuals and the corresponding error variances are determined by our procedure of estimation. Had we used some other procedure (such as the orthodox covariance technique) we would have been led to different, but appropriate, values of these residuals.

With van der Reyden's correction (ii) the residuals will be $\frac{1}{2}(\epsilon_u + \epsilon_w)$, and the corresponding error variance given by the analysis of variance of the corrected values will be

$$\tfrac{1}{4}\{V(\epsilon_u) + V(\epsilon_w) + 2 \text{ cov } (\epsilon_u , \epsilon_w)\} \tag{2}$$

Obviously expressions (1) and (2) will not in general be equivalent, and consequently the analysis of variance of the corrected values does not give a correct estimate of the error to which these values are subject.

Inspection of van der Reyden's Table II lends support to the view that his value for the error variance of the corrected values is a serious underestimate. If cov $(\epsilon_u , \epsilon_w)$ is zero the correct error variance would be four times that given by van der Reyden's method. In this case van der Reyden's value would be one quarter of the sum of the control and treated error variances, namely 13.0. The value actually obtained by van der Reyden, 9.4, thus suggests that cov $(\epsilon_u , \epsilon_w)$ has a small negative value.

A number of effects and interactions have been judged highly significant for the corrected data. There seems little scientific or agricultural reason why some of these should be so. For instance, the interactions RT and $R'T$ are each highly significant, whereas neither R nor R' reach significance at the 5% level. Yates [3] has pointed out that unexpected significance of interactions may be a warning of faulty analytical procedure.

Further, inspection of van der Reyden's Table I gives little indication of the trends one would expect if the treatment differences really were significant. For instance, there are no very noticeable trends associated with increasing concentration of chemicals, or with time of application.

As already mentioned, there is some evidence that correction (i) has been incorrectly applied to the numerical example. An inspection of total stand losses for two replicates as given by van der Reyden indicates that the correction has been added rather than subtracted. It will be noted from his Table I that when stand losses are high on controls the correction is nearly always positive, the opposite occuring when control stand losses are low. Indeed, if the entries are arranged in descending order of stand losses for control plots, in the first 21 cases the correction results in an increased figure in all but two cases. For the remaining 21 cases the correction is zero or negative in all but two cases. The net effect is to give a greater spread of stand losses, resulting in an increase of both treatment and total sums of squares, as is evident from van der Reyden's Table II.

A further point of criticism of van der Reyden's procedure arises from his statement that if "treatment" effects on the control plots do not prove to be significant, no further attention need be given to the data from them. This recommendation seems to illustrate a misconception.

The non-random distribution of insects over the experimental area only represents in an aggravated form the same problem as arises from the non-random distribution of soil fertility in a fertilizer experiment or variety trial. The use of randomisation, as explained by Fisher [1] is designed to overcome this difficulty. If randomisation has been correctly carried out the straightforward analysis of the treated plots will be *valid*, even though the *precision* of the experiment may not be high. Furthermore, if correctly randomised, the "treatment" effects in an analysis of the control plots will only prove significant in 5% of all trials on the average, if the 5% significance level is chosen, unless there has been some carry over of treatment effects from neighboring plots.

It may well be possible to improve the precision of the experiment by using the results from the control plots. In general, the appropriate statistical technique is the analysis of covariance. Van der Reyden's objection to this, that "the more effective a treatment, the less the correlation between treated and control plots" would appear to have relevance only in the case of highly effective treatments which reduce stand losses to zero. What matters is the correlation between the residuals on the treated and control plots, and except in the case just mentioned this may well be high regardless of treatment effects.

Whereas van der Reyden recommends using control results only when "treatment" affects are significant on control plots, in this case the very fact that they are significant should be taken as a warning

to proceed with caution, in case, as already suggested, the treatments have affected control plot observations by a carry over effect. Before control plot observations, or other sources, are used to provide supplementary information the experimenter must satisfy himself that this information does not itself reflect treatment effects. For instance, in field trials, counts of the numbers of seeds germinating, or of numbers of plants maturing, should not be used as supplementary information if these numbers are themselves influenced by the treatments under investigation. On the other hand, in a field trial it would be perfectly safe to use as supplementary information the yields of a uniformity trial conducted on the same land in the year previous to that in which the treatments under test were applied, since the treatments cannot have affected these yields.

In conclusion, it may be remarked that the irregular distribution of insects in experiments such as that described by van der Reyden raises important problems of technique if useful information is to be obtained. This writer does not know if it was practicable in this case to use insect counts on soil samples to obtain supplementary information. This technique has been widely used.

Acknowledgements

This note was prepared whilst the writer was on study leave at Rothamsted Experimental Station. He is indebted to Dr. F. Yates and Mr. M. J. R. Healy for useful advice and guidance during its preparation.

REFERENCES

1. Fisher, R. A. *The Design of Experiments.* Oliver & Boyd, 6th Edn., 1951.
2. van der Reyden, D. Design and Analysis of Soil Insecticide Field Experiments. *Biometrics, 10,* 291, 1954.
3. Yates, F. The Principles of Orthogonality and Confounding in Replicated Experiments. *J. Agri. Sci., 23,* 108, 1933.

COVARIANCE ANALYSIS AS AN ALTERNATIVE TO STRATIFICATION IN THE CONTROL OF GRADIENTS

ANNE D. OUTHWAITE AND A. RUTHERFORD*

Department of Statistics, University of Aberdeen

Introduction

In a recent paper, Federer and Schlottfeldt (1) illustrated the use of covariance to control gradients in an experiment as a substitute for deliberate stratification in the design. For this purpose, they took account of linear and quadratic trends. Since there was no obvious reason for stopping at this stage, we have examined the effect of including all terms up to the sixth degree.

The Covariance Analysis

Federer and Schlottfeldt discussed measurements of the heights of tobacco plants in an experiment on seven treatments arranged in eight randomized blocks. A fertility gradient within the blocks was suspected and they therefore calculated a quadratic covariance analysis on a measure of distance in this direction. For the study of a regression trend of higher degree, the computations are simplified by using standard orthogonal polynomial values from Fisher and Yates's Statistical Tables (2) based upon distance from the centre of the experiment, this modification involving no difference in principle. Table I reproduces the yields from (1) and also the covariates, x_1 to x_6 for the corresponding orthogonal polynomials.

The analysis of squares and products up to the third degree is shown in Table II, which includes the quantities required for subsequent covariance adjustments and agrees with Tables III and IV of (1) with respect to x_1 , x_2 and y.

To estimate the regression coefficients in the cubic analysis, the following set of equations must be solved.

$$\begin{aligned} 214.750b_1 + \ 9.500b_2 - \ 5.375b_3 &= \ 4{,}559.7 \\ 9.500b_1 + 585.250b_2 - \ 7.625b_3 &= \ 17{,}906.6 \\ -5.375b_1 - \ 7.625b_2 + 38.000b_3 &= -3{,}238.6 \end{aligned}$$

*On study leave from Ministry of Agriculture, N. Ireland.

TABLE I

Plant Height, First Measurement (Total for 20 Plants in cms)

Covariate						Replicate Number								Totals
x_1	x_2	x_3	x_4	x_5	x_6	1	2	3	4	5	6	7	8	
−3	5	−1	3	−1	−1	*F* 1299.2	*B* 1369.2	*A* 1169.5	*F* 1219.1	*F* 1120.0	*G* 1031.5	*B* 1076.4	*D* 1099.6	9384.5
−2	0	1	−7	4	6	*G* 875.9	*E* 844.2	*E* 975.8	*A* 971.7	*G* 827.0	*B* 846.5	*A* 917.9	*E* 947.4	7206.4
−1	−3	1	1	−5	−15	*D* 960.7	*F* 968.7	*C* 873.4	*G* 607.6	*D* 671.9	*D* 667.8	*E* 627.6	*B* 787.1	6164.8
0	−4	0	6	0	20	*C* 1004.0	*G* 975.5	*G* 797.8	*D* 1000.0	*C* 972.2	*C* 853.6	*F* 776.4	*A* 898.3	7277.8
1	−3	−1	1	5	−15	*A* 1173.2	*C* 1322.4	*B* 1069.7	*B* 1343.3	*A* 1083.7	*A* 1087.1	*C* 960.4	*C* 1174.9	9214.7
2	0	−1	−7	−4	6	*B* 1031.9	*A* 1172.6	*F* 1093.3	*E* 999.4	*B* 1146.9	*E* 990.2	*G* 852.4	*F* 1003.3	8290.0
3	5	1	3	1	−1	*E* 1421.1	*D* 1418.9	*D* 1169.6	*C* 1181.3	*E* 993.8	*F* 1021.9	*D* 1006.2	*G* 947.6	9160.4
0	0	0	0	0	0	7766.0	8071.5	7149.1	7322.4	6815.5	6498.6	6217.3	6858.2	56698.6

TABLE II

Cubic Covariance Analysis

Source	D.F.	$\sum x_1^2$	$\sum x_2^2$	$\sum x_3^2$	$\sum x_1x_2$	$\sum x_1x_3$	$\sum x_2x_3$	$\sum x_1y$	$\sum x_2y$	$\sum x_3y$	$\sum y^2$
Total	55	224.00	672.00	48.00	0.000	0.000	0.000	4544.8	17414.8	−4357.6	1,931,777
Replicates	7	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	388,314
Treatments	6	9.25	86.75	10.00	−9.500	5.375	7.625	−14.9	−431.8	−1119.0	273,875
Error	42	214.75	585.25	38.00	9.500	−5.375	−7.625	4559.7	17906.6	−3238.6	1,269,588
$E + T$	48	224.00	672.00	48.00	0	0	0	4544.8	17474.8	−4357.6	1,543,463

TABLE III
Regression Analysis

Degree of Polynomial	D.F.	Sum of Squares removed by Regression Analysis					
		First Degree	Second Degree	Third Degree	Fourth Degree	Fifth Degree	Sixth Degree
First	1	96816					
Second	1		535991				
Third	1			222767			
Fourth	1				17088		
Fifth	1					130116	
Sixth	1						17385
D.F. for Error S.S.	42	41	40	39	38	37	36
Error Sum of Squares	1269588	1172772	636781	414014	396926	266810	249425
Error Mean Square	30228	28604	15920	10616	10445	7211	6928
F		3.38	33.67	20.99	1.63	18.04	2.51

As the variances of the adjusted means were required, it was simplest to invert the matrix of the coefficients giving:—

$$(c_{ij}) = \begin{bmatrix} 0.0046758 & -0.0000675 & 0.0006478 \\ -0.0000675 & 0.0017141 & 0.0003344 \\ 0.0006478 & 0.0003344 & 0.0264745 \end{bmatrix}$$

and so:—

$$b_1 = 18.01396$$

$$b_2 = 29.30342$$

$$b_3 = -76.79897$$

The same process was followed for the regression coefficients in the higher degree covariance analyses. Table III shows the amount by which the error sum of squares was reduced by successive steps in the analysis up to the sixth degree term. The first entry in each column is obtained by subtracting the error sum of squares for that term from the previous error sum of squares. The error sum of squares for the quadratic differs from that in (1) because of an arithmetical error in the original paper. Tests of significance made by comparing the square for each term of the regression with the corresponding error mean square are open to criticism in that successive tests are not independent, but they strongly indicate that a covariance adjustment ought to include the third and fifth degree components though the fourth and sixth are of little importance for this set of data.

The Adjusted Means

As orthogonal polynomials were used, each $\bar{x}_i$ was zero. Hence the formulae for the adjusted treatment means $\bar{Y}'_i$, (where the $\bar{Y}_i$ are the unadjusted treatment means) are

$$\bar{Y}'_i = \bar{Y}_i - \sum_{j=1}^{n} b_j \bar{x}_{ji}$$

for the nth degree covariance where the b_j are the regression coefficients calculated for each covariance analysis, as given in Table IV.

From these, the means in Table V are obtained; these have been given only to one decimal place, as there appears to be little object in going further, when the standard errors are over 30 even after the sixth degree covariance adjustment.

TABLE IV
Regression Coefficients

Degree of Covariance Analysis	b_1	b_2	b_3	b_4	b_5	b_6
First	21.22934					
Second	19.89326	30.27350				
Third	18.01396	29.30342	−76.79897			
Fourth	18.40304	28.94409	−76.48360	4.21134		
Fifth	18.65597	26.84553	−80.88729	3.81832	15.34852	
Sixth	18.75972	26.80314	−81.34208	3.94426	15.32957	−1.58610

Variance of Adjusted Means

The general formula for the variance of the difference between two means adjusted for linear regression is

$$V(\bar{Y}'_1 - \bar{Y}'_2) = s^2\left(\frac{2}{r} + \frac{(\bar{x}_{1i} - \bar{x}_{2i})^2}{A}\right)$$

where s^2 is the residual error mean square for y after the removal of the regression component, r is the number of replicates, and A is the error sum of squares for x. This variance depends on the pair of treatments compared; Finney (3) pointed out that, if treatment differences in x are fairly small, this inconvenience could be avoided by averaging the second term over all possible pairs. He showed this to be equivalent to taking the variance of any one adjusted mean as

$$V(\bar{Y}'_i) = \frac{s^2}{r}\left(1 + \frac{d}{A(t-1)}\right)$$

where d is the sum of squares for "treatments" for x, and t is the number of treatments. By generalising this the following formula is found for nth degree polynomial:

$$V(\bar{Y}'_i) = \frac{s^2}{r}\left(1 + \frac{1}{(t-1)}\sum_{i,j=1}^{n} c_{ij}d_{ij}\right)$$

where the d_{ij} is the "treatment" sum of products of x_i and x_j, and the c_{ij} are the elements of the inverse matrix of the error sums of squares and sums of products.

Gains in Information

Federer and Schlottfeldt (1) compared the information obtained from alternative covariance analyses essentially in terms of s^2, without

TABLE V

Unadjusted Treatment Totals and Adjusted Treatment Means of Plant Height, for Second, Third and Fifth Degree

Treatments	Treatment Totals						Means for 20 Plants (cms)		
	x_{1i}	x_{2i}	x_{3i}	x_{4i}	x_{5i}	Y_i	Second Degree Adjustment	Third Degree Adjustment	Fifth Degree Adjustment
A	−2	−8	−3	−9	18	8474.0	1094.5	1064.3	1030.2
B	−3	1	−4	−12	−1	8671.0	1087.6	1048.6	1054.7
C	5	−19	−1	25	11	8342.2	1102.2	1091.5	1051.7
D	3	7	5	21	−13	7994.7	965.4	1014.9	1034.3
E	3	7	4	−28	1	7799.5	941.0	980.9	996.3
F	−3	13	−3	5	−15	8501.9	1021.0	993.1	1022.2
G	−3	−1	2	−2	−1	6915.3	875.7	894.0	897.9
Standard Errors							±45.3	±37.8	±32.4

allowance for the errors of estimation of regression coefficients. Table VI illustrates the consistent underestimation of variances to which this leads, the omitted terms being always positive. We consider that the most easily interpreted measure of gain in information is given by a comparison of variances of treatment means with different covariance adjustments, always making use of the formula of the last section. Table VII shows these relative efficiencies and also, to compare with the original paper, the units of information assessed on a basis of a standard error equal to 5% of the general mean.

TABLE VI

Variance of a Treatment Adjusted for Covariance With and Without Corrections

Degree of Covariance Adjustment	Without Corrections	With Corrections
Unadjusted	3778.5	3778.5
First	3575.5	3601.2
Second	1990.0	2053.2
Third	1327.0	1431.0
Fourth	1305.6	1470.3
Fifth	901.4	1052.4
Sixth	866.0	1021.4

TABLE VII

Comparison With Unadjusted Variance and Gains in Information

Degree of Covariance Adjustment	Comparisons With	
	Unadjusted Variance of Y	A Precision Equal to 5% of Mean
Unadjusted	1.00	0.65
First	1.05	0.68
Second	1.84	1.19
Third	2.64	1.71
Fourth	2.57	1.66
Fifth	3.59	2.31
Sixth	3.70	2.38

Comparison with Latin Square

If an experimental area were suspected of having fertility trends in two directions, the natural design to adopt would be a Latin square or

some modification of this. It is therefore natural to enquire how a 7×7 Latin square would have compared with the eight randomized blocks actually used in this experiment. Our analysis may be regarded as an attempt to obtain some of the advantages of a Latin square from an experiment in randomized blocks. Indeed, if a Latin square design were first analysed by calculating sums of squares for treatments and for columns but not for rows, and then a multiple covariance were used to eliminate polynomial trends of the highest possible degree between rows, the outcome would be exactly the same as that of the ordinary Latin square analysis; the calculations described above would become trivial. The orthogonality property of the Latin square makes valid the much more convenient procedure of direct calculation of a sum of squares between rows.

The residual error mean square obtained in our analysis, after eliminating all polynomial trends up to the sixth degree, estimates the mean square that would have been obtained if a 7×7 Latin square had been used on the same area with the same size of plot (of course omitting seven of the plots, or one block of the actual experiment). Consequently the error variance of a treatment mean for such a Latin square experiment may be estimated to be 989.7, which may be compared with the final figure obtained for the average variance of a treatment mean in the randomized block design 1021.4. Despite the additional replication, the randomized block design seems to compare the treatments with slightly less precision (at least in respect of this one measurement) than a Latin square design would have done. The corrections to the variance in the randomized block experiment, arising from sampling errors in the estimation of regression coefficients, more than counterbalance the gain of one replicate. The randomized block design has the further advantage of estimating the error with 36 degrees of freedom instead of 30 from the Latin square, but, when the number of degrees of freedom is of this order, such an increase is scarcely worth having.

On the evidence of this experiment then, nothing has been gained by having one replicate more than the Latin square would give, as the increase in precision has been lost by the destruction of orthogonality. This finding for the analysis of one measurement in one experiment does not establish any general principle; it does suggest that experimenters should be cautious in departing from a simple orthogonal design merely to achieve a slight increase in replication.

In the circumstances of this experiment, a compromise might have been effected by constructing the design as a 7×7 Latin square with one additional column. Cochran and Cox (4) have discussed such designs in their Chapter 13, and have shown that the design with seven

treatments and eight replicates has 98% efficiency; their measure of efficiency can be shown to be equal to the factor

$$\frac{(t-1)}{(t-1)+\sum c_{ij}d_{ij}}$$

In the design used, however, the corresponding efficiency was only 85%. Thus as compared with the 7×7 Latin square the extra replication in the Cochran and Cox design would give a net gain in precision of 12% ($0.98 \times 8/7 = 1.12$), while for the design used the corresponding figure was a net loss of 3% ($0.85 \times 8/7 = 0.97$)

Summary

Federer and Schlottfeldt (1) discussed the analysis of a randomized block design with eight blocks and seven treatments in which there was a fertility trend within the blocks. They used a covariance analysis to eliminate linear and quadratic components of this trend. We have extended this to the limit, using orthogonal polynomials to the sixth degree. We have pointed out that the variances of the adjusted means should be corrected to allow for the sampling errors of the regression coefficients, and have given a general formula for the average value of the corrections. These corrected variances were used to calculate the gains in information from the covariance analysis. Even after the full covariance adjustment, the design used turns out to be less efficient than a 7×7 Latin square with the same variance per plot, as the advantage gained by having an extra replicate was lost by the extreme non-orthogonality. For similar circumstances, Cochran and Cox (4) have suggested a design with eight replicates consisting of a 7×7 Latin square with an extra column; this design would have been a better choice, since it departs so little from orthogonality that the gain in using the extra replicate outweighs the loss due to nonorthogonality and leaves an overall gain in precision of 12% (assuming equal variances per plot).

REFERENCES

(1) Federer, W. T. & Schlottfeldt, C. S. (1954): The use of covariance to control gradients in experiments. *Biometrics, 10,* 282–290.

(2) Fisher, R. A. & Yates, F. (1953): *Statistical Tables for Biological, Agricultural and Medical Research.* (3rd edition). London: Oliver & Boyd.

(3) Finney, D. J. (1946): Standard errors of yields adjusted for regression on an independent measurement. *Biometrics, 2,* 53–55.

(4) Cochran, W. G. & Cox, G. M. (1950): *Experimental Designs.* New York: John Wiley & Sons.

ON THE ANALYSIS OF VARIANCE OF A TWO-WAY CLASSIFICATION WITH UNEQUAL SUB-CLASS NUMBERS

CLYDE YOUNG KRAMER
Virginia Agricultural Experiment Station
Blacksburg, Virginia

In many avenues of research it is necessary to analyse the variance of data which are classified in two ways with unequal numbers of observations falling into each sub-class of the classification. For data of this kind special methods of analysis are required because the inequality of the sub-class numbers causes lack of orthogonality among the main effects and interaction comparisons.

Table I below gives the basic notation for dealing with an analysis of a two-way classification with unequal sub-class numbers.

Several writers have dealt with the analysis of data of this form and various methods have been put forward. Some of the more prominent articles and discussions are cited below [1–13].

A simple preliminary step common to all methods is to separate the variance within sub-classes from the variance between sub-classes. Table II gives the analysis of variance for this preliminary step.

The problem of extending the analysis to the main effects and to the interaction between the main effects now arises. The $(pq - 1)$ degrees of freedom for between sub-classes can be partitioned in the usual way into $(p - 1)$ degrees of freedom for between A classes, $(q - 1)$ degrees of freedom for between B classes and $(p - 1)(q - 1)$ degrees of freedom for the interaction between the two classifications. The main difficulties arise in determining the correct sums of squares to be associated with each of these. One difficulty is that the addition theorem for sums of squares does not apply unless the sub-class numbers are proportional, and thus the interaction sum of squares cannot be computed by the usual method of differences. In fact, situations may occur where this procedure would give a negative result for the sum of squares for interaction.

Frequently we assume, from the nature of the data or from previous information or experience, that interaction is absent or if present, negligible. Making this assumption, we are interested in testing if there are any significant differences between the A-classes and between the B-classes.

TABLE I
Basic Notation

A Classes		B Classes						
		B_1	B_2	$\cdots$	B_j	$\cdots$	B_q	Total
A_1	No.	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1q}	$n_{1.}$
	Total	Y_{11}	Y_{12}	$\cdots$	Y_{1j}	$\cdots$	Y_{1q}	$Y_{1.}$
	Mean	y_{11}	y_{12}	$\cdots$	y_{1j}	$\cdots$	y_{1q}	$y_{1.}$
A_2	No.	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2q}	$n_{2.}$
	Total	Y_{21}	Y_{22}	$\cdots$	Y_{2j}	$\cdots$	Y_{2q}	$Y_{2.}$
	Mean	y_{21}	y_{22}	$\cdots$	y_{2j}	$\cdots$	y_{2q}	$y_{2.}$
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
A_i	No.	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iq}	$n_{i.}$
	Total	Y_{i1}	Y_{i2}	$\cdots$	Y_{ij}	$\cdots$	Y_{iq}	$Y_{i.}$
	Mean	y_{i1}	y_{i2}	$\cdots$	y_{ij}	$\cdots$	y_{iq}	$y_{i.}$
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
A_p	No.	n_{p1}	n_{p2}	$\cdots$	n_{pj}	$\cdots$	n_{pq}	$n_{p.}$
	Total	Y_{p1}	Y_{p2}	$\cdots$	Y_{pj}	$\cdots$	Y_{pq}	$Y_{p.}$
	Mean	y_{p1}	y_{p2}	$\cdots$	y_{pj}	$\cdots$	y_{pq}	$y_{p.}$
Total	No.	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.q}$	$n_{..}$
	Total	$Y_{.1}$	$Y_{.2}$	$\cdots$	$Y_{.j}$	$\cdots$	$Y_{.q}$	$Y_{..}$
	Mean	$y_{.1}$	$y_{.2}$	$\cdots$	$y_{.j}$	$\cdots$	$y_{.q}$	$y_{..}$

where: n_{ij} is the number of observations in the A_iB_j sub-class, $Y_{ij} = \sum_{k=1}^{n_{ij}} y_{ijk}$, y_{ijk} is the kth observation in the ijth sub-class,

$$y_{ij} = Y_{ij}/n_{ij}\ , \quad Y_{i.} = \sum_{j=1}^{q} Y_{ij}\ , \quad Y_{.j} = \sum_{i=1}^{p} Y_{ij}\ , \quad n_{i.} = \sum_{j=1}^{q} n_{ij}$$

$$n_{.j} = \sum_{i=1}^{p} n_{ij}\ , \qquad y_{i.} = Y_{i.}/n_{i.}\ , \qquad y_{.j} = Y_{.j}/n_{.j}\ ,$$

$$Y_{..} = \sum_{i=1}^{p} Y_{i.} = \sum_{j=1}^{q} Y_{.j}\ , \quad n_{..} = \sum_{i=1}^{p} n_{i.} = \sum_{j=1}^{q} n_{.j}\ , \quad y_{..} = Y_{..}/n_{..}\ .$$

TABLE II

Preliminary Analysis of Variance

Source	d.f.	S.S.	M.S.	F
Between Sub-Classes	$pq - 1$	$\sum_{i}^{p} \sum_{j}^{q} \frac{Y_{ij}^2}{n_{ij}} - \frac{Y_{..}^2}{n_{..}}$	s_B^2	s_B^2/s_w^2
Within Sub-Classes	$n_{..} - pq$	Subtraction	s_w^2	
Total	$n_{..} - 1$	$\sum_{i}^{p} \sum_{j}^{q} \sum_{k}^{n_{ij}} y_{ijk}^2 - \frac{Y_{..}^2}{n_{..}}$		

The optimum method of analysis of data of this type is the method of fitting constants from Yates [12]. Under the assumption of no interaction, a set of constants is fitted to the data so that the constants determine a set of sub-class means, with the property that the sum of weighted squares of the deviations of these means from the observed means is a minimum. If the classifications are large, this method becomes very tedious and laborious. In fact, one must write and solve at least $p + q$ normal equations, depending on which computational method is used.

Kendall [6] suggests using a much simpler and shorter, but less powerful test to compare the differences between the main effects. This method is known as the method of weighted squares of means. This incidentally, happens to be the optimum method if interaction is present. In this method unweighted marginal means are obtained.

$$(1) \qquad y'_{i.} = \frac{1}{q} \sum_{j=1}^{q} y_{ij} , \qquad y'_{.j} = \frac{1}{p} \sum_{i=1}^{p} y_{ij}$$

These are unbiased but inefficient estimates of the class means. By giving equal weight to all sub-class means, these marginal means for the A-classes are independent of the B-classes, and vice versa. The sums of squares due to the A-classes and B-classes are then calculated by (2).

$$(2) \qquad SS_A = \sum_{i=1}^{p} W'_{i.} y'^2_{i.} - \Big(\sum_{i=1}^{p} W'_{i.} y'_{i.} \Big)^2 \Big/ \sum_{i=1}^{p} W'_{i.}$$

$$SS_B = \sum_{j=1}^{q} W'_{.j} y'^2_{.j} - \Big(\sum_{j=1}^{q} W'_{.j} y'_{.j} \Big)^2 \Big/ \sum_{j=1}^{q} W'_{.j} ,$$

where $W'_{i.}$ and $W'_{.j}$ are given by (3).

$$(3) \qquad W'_{i.} = q^2 / \sum_{j=1}^{q} (1/n_{ij}), \qquad W'_{.j} = p^2 / \sum_{i=1}^{p} (1/n_{ij}).$$

$W'_{i.}$ and $W'_{.j}$ are the reciprocals of the variances (except for the factor σ^2) of $y'_{i.}$ and $y'_{.j}$. The mean squares of the A-classes and the B-classes are then tested with an F-test using the within sub-class mean square from the preliminary analysis. These are valid F tests but are low in power due to the inefficiency of the estimates $y'_{i.}$ and $y'_{.j}$ on which they are based.

A new method for this type of an analysis that is equally as simple to calculate as the method of weighted squares of means, but will be generally more powerful, will now be suggested.

For this new method, marginal means are obtained by (4).

$$(4) \qquad y''_{i.} = \frac{1}{n_{..}} \sum_{j=1}^{q} n_{.j} y_{ij}$$

$$y''_{.j} = \frac{1}{n_{..}} \sum_{i=1}^{p} n_{i.} y_{ij}$$

Since the weights $n_{.j}/n_{..}$ used in getting any mean, $y''_{i.}$, are independent of the row i concerned, the variation between the means $y''_{i.}$ is independent of the B-effects and vice versa as in the method of weighted squares of means.

This independence can be seen by considering the difference between any two A-class means, say $y''_{1.}$ and $y''_{2.}$. Since any $y_{ij} = m + a_i + b_j$, the marginal means are

$$(5) \qquad y''_{1.} = \frac{1}{n_{..}} \sum_{j=1}^{q} n_{.j}(m + a_1 + b_j)$$

$$= m + a_1 + \frac{1}{n_{..}} \sum_{j=1}^{q} b_j .$$

$$y''_{2.} = \frac{1}{n_{..}} \sum_{j=1}^{q} n_{.j}(m + a_2 + b_j)$$

$$= m + a_2 + \frac{1}{n_{..}} \sum_{j=1}^{q} b_j .$$

Now $y''_{1.} - y''_{2.} = a_1 - a_2$, which shows that the variation between the means $y''_{i.}$ is independent of the B-effects.

It will be seen that the derivations of the marginal means $y''_{i.}$ and $y''_{.j}$ tend to give weight to sub-class means more in proportion to the numbers on which they are based than do the corresponding derivations of $y'_{i.}$ and $y'_{.j}$ in the method of weighted squares of means.

The sum of squares due to the A-classes and B-classes for the new proposed method are then obtained by (6).

$$(6) \qquad SS_A = \sum_{i=1}^{p} W''_{i.} y''^2_{i.} - \Big(\sum_{i=1}^{p} W''_{i.} y''_{i.} \Big)^2 \Big/ \sum_{i=1}^{p} W''_{i.}$$

$$SS_B = \sum_{j=1}^{q} W''_{.j} y''^2_{.j} - \Big(\sum_{j=1}^{q} W''_{.j} y''_{.j} \Big)^2 \Big/ \sum_{j=1}^{q} W''_{.j} ,$$

where $W''_{i.}$ and $W''_{.j}$ are given by (7).

$$(7) \qquad W''_{i.} = n^2_{..} \Big/ \Big(\sum_{j=1}^{q} n^2_{.j}/n_{ij} \Big), \qquad W''_{.j} = n^2_{..} \Big/ \Big(\sum_{i=1}^{p} n^2_{i.}/n_{ij} \Big).$$

As in the method of the weighted squares of means, the weighting factors are the reciprocals of the variances (except for the factor σ^2) of the marginal means. The mean squares of the A-classes and the B-classes are then tested with an F-test using the within sub-class mean square from the preliminary analysis. These are also valid F-tests, but they are based on estimates of the class means, which will be generally more efficient and therefore more powerful than the F-tests obtained by the method of weighted squares of means.

This new method is proposed for cases in which p and q are large and when interaction can be assumed absent or negligible. If one is dubious about using the new method instead of the method of weighted squares of means, the following inequalities can be evaluated.

$$(8) \qquad \sum_{i=1}^{p} W''_{i.} - \sum_{i=1}^{p} W''^2_{i.} \Big/ \sum_{i=1}^{p} W''_{i.} > \sum_{i}^{p} W'_{i.} - \sum_{i=1}^{p} W'^2_{i.} \Big/ \sum_{i=1}^{p} W'_{i.}$$

$$\sum_{j=1}^{q} W''_{.j} - \sum_{j=1}^{q} W''^2_{.j} \Big/ \sum_{j=1}^{q} W''_{.j} > \sum_{j=1}^{q} W'_{.j} - \sum_{j=1}^{q} W'^2_{.j} \Big/ \sum_{j=1}^{q} W'_{.j}$$

If these inequalities are satisfied the new method will be more powerful than the weighted squares of means.

The above inequalities which compare the method of weighted squares of means and the new proposed method have been obtained in the following way. Starting with the model $y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$ we may write $y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$ where $\epsilon_{ij} = \sum_k \epsilon_{ijk}/n_{ij}$.

Dealing first with the method of weighted squares of means we have

$$y'_{i.} = \sum_j y_{ij}/q \tag{9}$$

$$= \mu + \alpha_i + \sum_j \beta_j/q + \sum_j \epsilon_{ij}/q,$$

which reduces to

$$y'_{i.} = \mu + \alpha_i + \epsilon'_{i.}\,, \tag{10}$$

where $\epsilon'_{i.} = \sum_j \epsilon_{ij}/q$ on applying the restriction $\sum_i \beta_i = 0$. Now

$$E(SS_A) = E \sum_i W'_{i.}(\alpha_i - \bar{\alpha} + \epsilon'_{i.} - \bar{\epsilon})^2, \tag{11}$$

where

$$\bar{\alpha} = \sum_i W'_{i.}\alpha_i / \sum_i W'_{i.} \quad \text{and} \quad \bar{\epsilon} = \sum_i W'_{i.}\epsilon'_{i.} / \sum_i W'_{i.}\,.$$

Since the cross product terms are zero and since $W'_{i.}$ is the reciprocal of the variance of $\epsilon'_{i.}$ (except for the factor σ^2) this reduces to

$$E(SS_A) = \sum_i W'_{i.}(\alpha_i - \bar{\alpha})^2 + (p-1)\sigma^2 \tag{12}$$

Turning now to the proposed new method

$$y''_{i.} = \sum_j n_{.j}y_{ij}/n_{..} \tag{13}$$

$$= \mu + \alpha_i + \sum_j n_{.j}\beta_j/n_{..} + \sum_j n_{.j}\epsilon_{ij}/n_{..}$$

which reduces to

$$y''_{i.} = \mu + \alpha_i + \sum_j n_{.j}\beta_j/n_{..} + \epsilon''_{i.}\,, \tag{14}$$

where

$$\epsilon''_{i.} = \sum_j n_{.j}\epsilon_{ij}/n_{..}\,.$$

Now

$$E(SS_A) = E \sum_i W''_{i.}(\alpha_i - \bar{\alpha}' + \epsilon''_{i.} - \bar{\epsilon}')^2, \tag{15}$$

where

$$\bar{\alpha}' = \sum_i W''_{i.}\alpha_i / \sum_i W''_{i.} \quad \text{and} \quad \bar{\epsilon}' = \sum_i W''_{i.}\epsilon''_{i.} / \sum_i W''_{i.}\,.$$

This reduces to

$$E(SS_A) = \sum_i W''_{i.}(\alpha_i - \bar{\alpha}')^2 + (p-1)\sigma^2. \tag{16}$$

From the expected values of the main effect sum of squares for the respective methods (12) and (16), it can be seen that the new method

will be more powerful for detecting real effects if

$$\sum_i W''_{i.}(\alpha_i - \bar{\alpha}')^2 > \sum_i W'_{i.}(\alpha_i - \bar{\alpha})^2. \tag{17}$$

In the absence of knowledge of the α_i values it seems reasonable to replace each α_i^2 by a constant α^2 and each product term $a_i a_j$, $i \neq j$ by zero. On doing this the inequality (17) is readily seen to reduce to (8).

TABLE III

Data for Milk Yields, in Pounds, of Cows Freshening in Two Seasons

			Seasons	
			Fall and Winter	Spring and Summer
COWS	1	No.	2	2
		Total	969	632
		Mean	484.50	316.00
	2	No.	7	1
		Total	2477	262
		Mean	349.57	262.00
	3	No.	2	2
		Total	827	540
		Mean	413.50	270.00
	7	No.	2	4
		Total	572	890
		Mean	286.00	222.50
	9	No.	5	3
		Total	880	539
		Mean	176.00	179.67
	11	No.	7	4
		Total	2703	1280
		Mean	386.14	320.00
	21	No.	3	3
		Total	1184	1194
		Mean	394.67	398.00
	27	No.	4	5
		Total	1753	1701
		Mean	438.25	340.20

The quantities in (8) are also the coefficients of σ_i^2 and σ_{ik}^2 in the expectations of the mean squares under Model II analysis of variance. Under the no interaction assumption and using the within sub-class mean square for testing, it is clear that the proposed method will be more sensitive if the coefficients of the "variance component" under test are larger, as was pointed out by a referee of this paper.

Numerical example: The following example was taken from pages 341–346, *Methods of Statistical Analysis* by C. H. Goulden [4].

Goulden gives the analysis of variance in Table IV for these data (Table III) by the method of fitting constants.

TABLE IV
Analysis of Variance (Fitting Constants)

Source	d.f.	S.S.	M.S.	F	5% Point
Cows	7	322,714	46,102	4.58	2.25
Seasons	1	58,177	58,177	5.78	4.08
Cows × Seasons	7	35,436	5,062		
Error	40	402,860	10,072		

It is seen that there is no evidence of an interaction, so the optimum method of completing the analysis would be by fitting constants which has been done.

Goulden also gives the analysis in Table V for these data by the method of weighted squares of means.

TABLE V
Analysis of Variance (Weighted Squares of Means)

Source	d.f.	S.S.	M.S.	F	5% Point
Cows	7	289,191	41,313	4.10	2.25
Seasons	1	64,810	64 810	6.43	4.08
Cows × Seasons	7	35,436	5,062		
Error	40	402,860	10,072		

It is seen that the method of weighted squares of means gives a sum of squares for cows 33,523 less than the method of fitting constants and a sum of squares for seasons 6,633 more than the method of fitting constants.

The weights used in the analysis are as follows:

$$W'_{1.} = 4.00000 \qquad W'_{6.} = 10.18184$$

$$W'_{2.} = 3.50000 \qquad W'_{7.} = 6.00000$$

$$W'_{3.} = 4.00000 \qquad W'_{8.} = 8.88888$$

$$W'_{4.} = 5.33332 \qquad W'_{.1} = 24.911936$$

$$W'_{5.} = 7.50000 \qquad W'_{.2} = 19.009920$$

Table VI gives the required calculations for obtaining the analysis of variance for the proposed new method.

Using the totals in Table VI we obtain the following sum of squares:

$$\text{Cows} = 5{,}776{,}003.795 - \frac{(16{,}533.135)^2}{49.99691} = 308{,}775$$

$$\text{Seasons} = 5{,}371{,}198.409 - \frac{(16{,}025.704)^2}{48.31119} = 55{,}180.$$

Comparing these sums of squares with the ones obtained by fitting constants, it is seen that for cows the new method is only 13,939 less and for seasons 2,997 less. Since there was no indication of an interaction, the proposed method gave results more near the results obtained by the method of fitting constants than did the method of weighted squares of means.

If we had used the inequalities to decide which method would be best before we did the analysis, we would have obtained the following values:

$$\text{Cows} \quad 42.84540 > 42.36764$$

$$\text{Seasons} \quad 22.86865 > 21.564385.$$

The above inequalities tell us that the new proposed method will be more powerful than the method of weighted squares of means for both cow and season effects.

If we have a $p \times q$ table where p and q are both greater than 2, the only way we can obtain a test of interaction is by the method of fitting constants. If one does not wish a test of interaction or can assume that it is negligible, one can be sure he is obtaining the best approximation to the method of fitting constants by calculating the inequalities

TABLE VI

CALCULATIONS FOR PROPOSED NEW METHOD

		Fall and Winter	Spring and Summer	$n_{i.}$	$y''_{i.} = (\sum_j n_{.j} y_{ij})/n_{..}$	$\sum_j n^2_{.j}/n_{ij}$	$W''_{i.} = n^2_{../}(\sum_j n_{.j}/n_{ij})$	$W''_{i.}y''_{i.}$	$W''_{i.}y''^2_{i.}$
C_1	No.	2	2	4					
	Mean	484.50	316.00		412.286		3.92000	1616.161	666.320.554
	$n^2_{.j}/n_{ij}$	512.000000	288.000000			800.000000			
	$n^2_{i.}/n_{ij}$	8.000000	8.000000						
C_2	No.	7	1	8					
	Mean	349.57	262.00		312.040		4.34177	1354.806	422.753.664
	$n^2_{.j}/n_{ij}$	146.285714	576.000000			722.285714			
	$n^2_{i.}/n_{ij}$	9.142857	64.000000						
C_3	No.	2	2	4					
	Mean	413.50	270.00		352.000		3.92000	1379.840	485.703.680
	$n^2_{.j}/n_{ij}$	512.000000	288.000000			800.000000			
	$n^2_{i.}/n_{ij}$	8.000000	8.000000						
C_7	No.	2	4	6					
	Mean	286.00	222.50		258.786		4.78049	1237.124	320.150.371
	$n^2_{.j}/n_{ij}$	512.000000	144.000000			656.000000			
	$n^2_{i.}/n_{ij}$	18.000000	9.000000						
C_9	No.	5	3	8					
	Mean	176.00	179.67		177.573		7.90322	1403.398	249.205.593
	$n^2_{.j}/n_i$	204.800000	192.000000			396.800000			

	$n^2_{.j}/n_{.j}$	204.800000	192.000000			396.800000			
C_{11}	No.	7	4	11					
	Mean	386.14	320.00		357.794		10.80315	3865.302	1,382,981.864
	$n^2_{.j}/n_{ij}$	146.285714	144.000000			290.285714			
	$n^2_{i.}/n_{ij}$	17.285714	30.250000						
C_{21}	No.	3	3	6					
	Mean	394.67	398.00		396.097		5.88000	2329.050	922,529.718
	$n^2_{.j}/n_{ij}$	341.333333	192.000000			533.333333			
	$n^2_{i.}/n_{ij}$	12.000000	12.000000						
C_{27}	No.	4	5	9					
	Mean	438.25	340.20		396.229		8.44828	3347.454	1,326,358.351
	$n^2_{.j}/n_{ij}$	256.000000	115.200000			371.200000			
	$n^2_{i.}/n_{ij}$	20.250000	16.200000						
	$n_{.j}$	32	24	56	Totals		49.99691	16.533.135	5.776.003.795
	$y''_{.j} = (\sum_i n_{i.} y_{ij})/n_{..}$	358.435	288.967						
	$\sum_i n^2_{i.}/n_{ij}$	105.478571	168.783333						
	$W''_{.j} = n^2_{..}/(\sum_i n^2_{i.}/n_{ij})$	29.73116	18.58003		48.31119				
	$W''_{.j}\ y''_{.j}$	10.656.688	5.369.016		16,025.704				
	$W''_{.j}\ y''^2_{.j}$	3.819.729.963	1,551,468.446		5,371,198.409				

given by (8) as was done above and then decide on the method of analysis.

ACKNOWLEDGEMENTS

This new method of analysis was suggested to me by Dr. D. B. Duncan. I should like to thank Dr. Duncan for his helpful comments and direction.

REFERENCES

1. Anderson, R. L. and Bancroft, T. A. *Statistical Theory in Research.* McGraw-Hill Book Co., Inc., New York, 1952.
2. Brandt, A. E. The Analysis of Variance in a $2 \times s$ Table With Disproportionate Frequencies. *J. A. S. A.*, *28*, 1933, pp. 164–173.
3. Crump, S. L. The Estimation of Variance Components in Analysis of Variance. *Biometrics Bulletin*, *2*, 1946, pp. 7–11.
4. Goulden, C. H. *Methods of Statistical Analysis.* John Wiley and Sons, Inc., New York, 1952.
5. Hazel, L. N. The Covariance Analysis of Multiple Classification Tables with Unequal Sub-Class Numbers. *Biometrics Bulletin*, *2*, 1946, pp. 21–25.
6. Kendall, M. G. *The Advanced Theory of Statistics.* Vol. II, Third Edition, Hafner Publishing Co., New York, 1951.
7. Rao, C. R. *Advanced Statistical Methods in Biometric Research.* John Wiley and Sons, Inc., New York, 1952.
8. Snedecor, G. W. *Statistical Methods.* The Iowa State College Press, Ames, Iowa, 1950.
9. Snedecor, G. W. and Cox, G. M. Disproportionate Sub-class Numbers in Tables of Multiple Classification. *Research Bulletin No. 180*, Agricultural Experiment Station, Iowa State College of Agriculture and Mechanic Arts, Ames, Iowa, 1935.
10. Stevens, W. L. Statistical Analysis of a Non-Orthogonal Tri-Factorial Experiment. *Biometrika*, *35*, 1948, pp. 346–367.
11. Wilks, S. S. *Mathematical Statistics.* Princeton University Press, Princeton, N. J., 1950.
12. Yates, F. The Principles of Orthogonality and Confounding in Replicated Experiments. *Journal of Agricultural Science*, *23*, 1933, pp. 108–145.
13. Yates, F. The Analysis of Multiple Classifications With Unequal Numbers in the Different Classes. *J. A. S. A.*, *29*, 1934, pp. 51–66.

A METHOD OF ANALYSIS FOR A DOUBLE CLASSIFICATION ARRANGED IN A TRIANGULAR TABLE

NELL DITCHBURNE

Division of Mathematical Statistics, C.S.I.R.O., Melbourne

I. INTRODUCTION

When an experiment is designed to test the effect of two factors each at several levels, on some measurable quantity, the data may be arranged in a two-way table. When each factor is tested at all levels of the other factor, and the number of observations in the subclasses are equal or proportional, estimates of the effects of each factor are easily obtained, the analysis of variance of the data is simple, and interpretation of results straightforward, as the data are orthogonal. However, when the numbers in the subclasses are unequal, or when all levels of one factor are not tested at all levels of the other, so that some subclasses are completely missing, the data are non-orthogonal, and a method of analysis must be found to suit the design of the particular experiment.

The method of analysis described in this paper is suitable for use with data which may be arranged in a triangular table such as that set out diagrammatically below. In the diagram, data are available for the subclasses marked x.

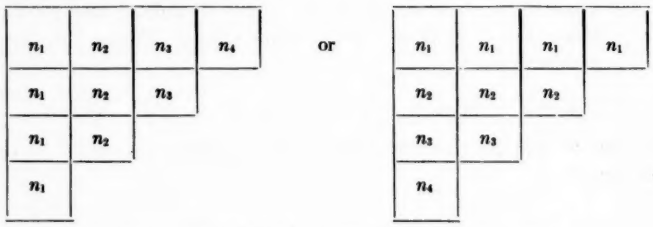
Rows	Columns			
	1	2	3	4
1	x	x	x	x
2	x	x	x	
3	x	x		
4	x			

The triangular type of design can often occur in experiments in which the levels of one factor may be found from the sum or difference of two other factors. For instance, in certain experiments in animal production, the effects of age at mating, year dropped, and year mated on percentage reproduction are investigated. In this case, age at mating is equal to the difference between year mated and year dropped. The data could be arranged in triangular tables, taking two factors at a time, that is, year mated and year dropped, age at mating and year dropped, or age at mating and year mated.

In chemical work, the effect of varying temperature range, with differing initial and final temperatures, may be analysed by this method; thus the effects of any two of range, initial temperature and final temperature may be determined. This triangular type of arrangement was employed in an experiment on the chemical retting of flax; in this case the effects of range and final temperature were considered, because rets were held longer at the final temperatures than at initial temperatures. A numerical example from this experiment is given in Section IV.

An analysis of a table of this type is possible only on the assumption that interactions do not exist, that is, that the effects of the two factors are additive. On this assumption, the method of fitting constants by least squares, as described by Yates (1933), is appropriate for all experiments with multiple classifications, whether or not there are empty subclasses. By this method of analysis, the correctness of the assumption of non-existent interactions may be tested; if there is evidence that interactions do exist, the data must be reanalysed, to examine separately the effects of each factor. For the two examples quoted above, the method of analysis to be described would be appropriate for those modes of classification of the data which gave effects which were additive.

Data arranged in a triangular table may be analysed without solving least squares equations, the sum of squares for each factor, freed from the effects of the other factor, being obtained from a set of orthogonal comparisons. This method, which is the one described in this paper, is more rapid than that involving the solution of least squares equations. However, it is applicable only under certain conditions: either the number of replicates in each subclass of the table should be equal, which is the general case discussed in Section III, or the number of replicates in each subclass within a single level of one of the factors should be equal. In the latter case, the table of subclass numbers would be of the form



II. ORTHOGONAL COMPONENTS OF A SUM OF SQUARES

Cochran and Cox (1950) give a summary of the conditions under which comparisons among k treatment totals T_i are orthogonal. If the number of replicates of the ith treatment is n_i, then the function

$$Z_w = l_{w1}T_1 + l_{w2}T_2 + \cdots l_{wk}T_k$$

is a comparison among the T_i provided

$$n_1l_{w1} + n_2l_{w2} + \cdots n_kl_{wk} = 0.$$

Two comparisons are orthogonal if

$$n_1l_{11}l_{21} + n_2l_{12}l_{22} + \cdots n_kl_{1k}l_{2k} = 0.$$

The quantity Z_w^2/D_w, where

$$D_w = n_1l_{w1}^2 + n_2l_{w2}^2 + \cdots + n_kl_{wk}^2$$

is a component of the sum of squares for treatments, and has one degree of freedom.

Among the k treatments, there are $k - 1$ comparisons Z_w which are orthogonal, and therefore it follows that

$$\sum (Z_w^2/D_w) = \sum (T_i^2/n_i) - (\sum T_i)^2/\sum n_i .$$

III. METHOD OF ANALYSIS FOR TRIANGULAR CLASSIFICATIONS

A description of the analysis for a triangular design when all the numbers in all subclasses are equal to n will now be given.

If k is the number of rows and columns of the table, there are $k - 1$ degrees of freedom in the sum of squares for the main effect of either rows or columns; each of these sums of squares may be subdivided into $k - 1$ orthogonal components.

If the totals for the ith row and column be R_i and C_i respectively, then it may be shown that the quantities

$$Y_i = (i - 1)R_i - \sum_{j=1}^{i-1} R_j + \sum_{j=k-i+2}^{k} C_j \qquad (i = 2, 3 \cdots k)$$

are mutually orthogonal comparisons among the rows.

The quantity Y_i is made up of $n(k - i + 1)$ values each multiplied by $i - 1$, less $n(i - 1)\ (k - i + 1)$ other values. Hence the variance of Y_i is proportional to

$$n[(i - 1)^2(k - i + 1) + (i - 1)(k - i + 1)] = ni(i - 1)(k - i + 1)$$

The sum of squares for rows freed from column effects is therefore given by

$$\frac{1}{n} \sum_{i=2}^{k} Y_i^2/i(i - 1)(k - i + 1).$$

The sum of squares for columns, freed from row effects, is obtained similarly.

The sum of squares for rows, column effects being ignored, is

$$\frac{1}{n}\left[\sum_{i=1}^{k} \left(R_i^2/(k - i + 1)\right) - \left(\sum_{i=1}^{k} R_i \right)^2 / \tfrac{1}{2}(k^2 + k) \right]$$

and that for columns, row effects being ignored, is obtained in a similar manner.

The interaction sum of squares has $\frac{1}{2}(k - 1)\ (k - 2)$ degrees of freedom, and is obtained from the total sum of squares between subclasses minus the sums of squares for rows with columns eliminated and columns ignoring rows.

Estimates of the means for each factor, freed from effects of the other, may be obtained in the following way.

If the estimates for the ith row are r_i, then

$$r_2 - r_1 = Y_2/n(k - 1),$$

$$2r_3 - r_2 - r_1 = Y_3/n(k - 2),$$

and generally

$$(i - 1)r_i - \sum_{j=1}^{i-1} r_j = Y_i/n(k - i + 1);$$

also

$$kr_1 + (k - 1)r_2 + (k - 2)r_3 + \cdots r_k = 0.$$

The solutions of these equations are

$$r_1 = -\frac{1}{nk(k + 1)} \sum_{j=2}^{k} \frac{Y_j(k + j)}{j(j - 1)}$$

$$r_2 = \frac{1}{n}\left[\frac{Y_2}{(k - 1)} - \frac{1}{k(k + 1)} \sum_{j=2}^{k} \frac{Y_j(k + j)}{j(j - 1)} \right]$$

etc.

Generally,

$$r_i = \frac{1}{n}\left[\frac{Y_i}{(k-i+1)(i-1)} + \sum_{j=2}^{i-1}\frac{Y_j}{j(k-j+1)(j-1)} - \frac{1}{k(k+1)}\sum_{j=2}^{k}\frac{Y_j(k+j)}{j(j-1)}\right]$$

$$= \frac{1}{n}\left[\frac{Y_i}{(k-i+1)(i-1)} + \frac{1}{k(k+1)}\left(\sum_{j=2}^{i-1}\frac{Y_j}{k-j+1} - \sum_{j=1}^{k}\frac{Y_j(k+j)}{j(j-1)}\right)\right]$$

The difference between two estimates r_i and r_h, $(i > h)$, is

$$\frac{1}{n}\left[\frac{Y_i}{(i-1)(k-i+1)} - \frac{Y_h}{h(k-h+1)} + \sum_{j=h+1}^{i-1}\frac{Y_j}{j(j-1)(k-j+1)}\right]$$

As the Y_i are orthogonal, the variance of $r_i - r_h$ is simply

$$\frac{1}{n^2}\left[\frac{V(Y_i)}{[(i-1)(k-i+1)]^2} + \frac{V(Y_h)}{[h(k-h+1)]^2} + \sum_{j=h+1}^{i-1}\frac{V(Y_j)}{[j(j-1)(k-j+1)]^2}\right]$$

Since the variance of Y_i is proportional to

$$ni(i-1)(k-i+1),$$

it follows that $V(r_i - r_h)$ is proportional to

$$\frac{1}{n}\left[\frac{i(i-1)(k-i+1)}{[(i-1)(k-i+1)]^2} + \frac{h(h-1)(k-h+1)}{[h(k-h+1)]^2} + \sum_{j=h+1}^{i-1}\frac{j(j-1)(k-j+1)}{j(k-j+1)(j-1)^2}\right]$$

$$= \frac{1}{n}\left[\frac{i}{(i-1)(k-i+1)} + \frac{h-1}{h(k-h+1)} + \sum_{j=h+1}^{i-1}\frac{1}{j(k-j+1)(j-1)}\right].$$

Estimates of column means c_i, and variances of differences between two estimates c_i and c_h may be obtained in a similar manner.

IV. NUMERICAL EXAMPLE

The data presented in Table 1 are values of buffer capacity after retting for 99 hours, for four varieties of flax, *A*, *B*, *C* and *D*,

TABLE 1

Individual Values of Buffer Capacity.

Temp.	Variety	Range				Totals for Ranges				
		0	4	8	12	0 to 12	0 to 8	0 to 4	0	
40	*A*	7.42	6.69	6.40	5.84					
	B	7.92	9.90	6.43	9.33					
	C	7.61	6.18	9.53	6.81					
	D	7.54	6.84	7.14	6.37					
	Total	30.49	29.61	29.50	28.35	117.95	89.60	60.10	30.49	
36	*A*	6.60	5.68	5.85	4.96					
	B	7.58	6.82	7.32	5.72					
	C	6.45	6.40	5.97	5.71					
	D	6.55	7.26	6.42	6.00					
	Total	27.18	26.16	25.56	22.39	101.29	78.90	53.34	27.18	
32	*A*	5.21	4.76	4.89						
	B	6.74	6.31	5.47						
	C	5.98	5.98	4.54						
	D	6.52	6.01	5.41						
	Total	24.45	23.06	20.31			67.82	47.51	24.45	
28	*A*	3.80	3.88							
	B	5.22	5.20							
	C	4.50	4.18							
	D	4.64	4.38							
	Total	18.16	17.64					35.80	18.16	

24	*A*	2.89								
	B	3.25								
	C	2.90								
	D	3.63								
	Total	12.67							12.67	
Totals (40 to 24)	*A*									74.87
	B									93.21
	C									82.74
	D									84.71
	Total									335.53
40 to 24	all	112.95								
40 to 28		100.28	96.47							
40 to 32		82.12	78.83	75.37						
40 to 36		57.67	55.77	55.06	50.74					

four temperature ranges, 0, 4, 8 and 12 degrees Centigrade, and five final temperatures, 24, 28, 32, 36 and 40 degrees Centigrade. It is seen that this example differs slightly from the general case described in Section III, but a similar method of analysis is used. Before analysis, data should be arranged so that the table is of the form shown in section I; otherwise the formulae of section III are meaningless.

TABLE 2

Buffer Capacity at 99 hours Retting.

Analysis of Variance

	D.F.	Sum of Squares		Mean Square
		Range (T elim.)	Temp. (R elim.)	
Varieties	3	12.1584	12.1584	4.0528**
Range	3	*4.6622*	4.3798	1.5541*
Final Temperature	4	84.0110	*84.2934*	21.0734**
$R \times T$	6	1.3715	1.3715	0.2286(n)
Error	39	20.1167	20.1167	0.5158
Total	55	122.3198	122.3198	

The analysis of variance is shown in Table 2. A detailed description of the methods used in computing the required sums of squares and estimates is given below, together with various short methods of computation, which are preferable to direct use of the formulae of section III.

Sums of Squares

Varieties. As all temperatures and ranges are equally represented in all varieties, the sum of squares is obtained direct from the variety totals.

$$(74.87^2 + 93.21^2 + 82.74^2 + 84.71^2)/14 - (335.53^2)/56 = 12.1584$$

Range (a) unadjusted for temperature effect.

This sum of squares is obtained by the method generally employed when subclass numbers are unequal:

$$(112.95^2/20 + 96.47^2/16 + 75.37^2/12 + 50.74^2/8 - 335.53^2/56) = 4.3798$$

(b) with temperature effect eliminated.

To simplify computation of the values $Y_{..}$, certain extra totals, shown in Table 1, are required. For instance, for comparison of ranges 0 and

4, a sum of squares with one degree of freedom is obtained as

$$(96.47 - 112.95 + 12.67)^2/(2 \times 16).$$

This is a direct comparison of ranges 0 and 4, summed for temperatures 40 to 28:

$$(96.47 - 100.28)^2/(2 \times 16) = (-3.81)^2/32$$
$$= 0.4536.$$

From the totals for ranges 0, 4 and 8, summed for temperatures 40 to 32, a sum of squares is obtained for the comparison of the mean for ranges 0 and 4 with the mean for range 8:

$$(2 \times 75.37 - 82.12 - 78.83)^2/(6 \times 12) = (-10.21)^2/72$$
$$= 1.4478.$$

From the totals for ranges 0 to 12, summed for temperatures 40 and 36, the sum of squares for the comparison of the mean for range 12 with the mean for ranges 0, 4 and 8 is

$$(3 \times 50.74 - 57.67 - 55.77 - 55.06)^2/(12 \times 8) = (-16.28)^2/96$$
$$= 2.7608.$$

By the addition of these separate squares, we obtain the value of 4.6622, with 3 degrees of freedom, which is shown in Table 2.

Temperature (a) unadjusted for range effect.

The sum of squares, with 4 degrees of freedom, is

$$(117.95^2/16 + 101.29^2/16 + 67.82^2/12 + 35.80^2/8 + 12.67^2/4$$
$$- 335.53^2/56) = 84.0110.$$

(b) with effect of range eliminated.

The difference between means for temperatures 40 and 36 is free of any range effect; the sum of squares, with 1 degree of freedom, is

$$(101.29 - 117.95)^2/(2 \times 16) = (-16.66)^2/32$$
$$= 8.6736.$$

For ranges 0 to 8, the sum of squares for the comparison of the mean for temperatures 40 and 36 with the mean for temperature 32 is

$$(2 \times 67.82 - 101.29 - 117.95 + 50.74)^2/(6 \times 12)$$
$$= (2 \times 67.82 - 89.60 - 78.90)^2/(6 \times 12)$$
$$= (-32.86)^2/72$$
$$= 14.9969.$$

For ranges 0 to 4, the sum of squares for the comparison of the mean for temperatures 40, 36 and 32 with the mean for temperature 28 is

$$(3 \times 35.80 - 60.10 - 53.34 - 47.51)^2/(12 \times 8) = (-53.55)^2/96$$
$$= 29.8709.$$

For range 0, the sum of squares for comparison of the mean of temperatures 40 to 28 with the mean for temperature 24 is

$$(4 \times 12.67 - 30.49 - 27.18 - 24.45 - 18.16)^2/(20 \times 4)$$
$$= (-49.60)^2/80$$
$$= 30.7520.$$

The total sum of squares for temperatures, with range effect eliminated, is therefore

$$8.6736 + 14.9969 + 29.8709 + 30.7520$$
$$= 84.2934, \text{ with 4 degrees of freedom.}$$

A check for these sums of squares is that

Range unadjusted + temperature adjusted
= Range adjusted + temperature unadjusted,
i.e., $4.3798 + 84.2934 = 4.6622 + 84.0110$.

The multipliers used in obtaining the orthogonal comparisons are summarized in Table 3; from the Table it is easily verified that the comparisons satisfy the conditions for orthogonality given in Section II.

Interaction of Range and Temperature

This sum of squares is obtained from the sum of squares for subclasses minus the sums of squares for temperature adjusted and range unadjusted, or temperature unadjusted and range adjusted.

$$(30.49^2 + 29.61^2 + \cdots 12.67^2)/4 - 335.53^2/56 - 84.2934 - 4.3798$$
$$\text{or} \qquad - 84.0110 - 4.6622.$$

The interaction of varieties with other effects is used as an estimate of error for testing the significance of the effects of range and temperature and their interaction, as it is considered that the four varieties are replications of the experiment.

The interaction effect of range and temperature is not significant; this confirms the original hypothesis, and indicates that the analysis by this method gives a valid test of the significance of the effects of range and temperature.

TABLE 3

Multipliers used in obtaining orthogonal comparisons.

Range Effect		n_i	l_{ti} Range				Divisor for Sum of Squares
			0	4	8	12	
Totals over temps.	40 to 28	16	−1	+1			32
	40 to 32	12	−1	−1	+2		72
	40 to 36	8	−1	−1	−1	+3	96
Temperature effect			Temperature				
			40	36	32	28	24
Totals over ranges	0	4	−1	−1	−1	−1	+4	80
	0 to 4	8	−1	−1	−1	+3		96
	0 to 8	12	−1	−1	+2			72
	0 to 12	16	−1	+1				32

Estimates

As this example differs from the general case, estimates are more easily obtained by solution of equations using the Y_i which were obtained in calculation of sums of squares than by adjustment of the general solution for the r_i given in Section III.

The equations to be solved are

$$\begin{aligned}
-t_{40} + t_{36} &= -16.66/16 \\
-t_{40} - t_{36} + 2t_{32} &= -32.86/12 \\
-t_{40} - t_{36} - t_{32} + 3t_{28} &= -53.55/8 \\
-t_{40} - t_{36} - t_{32} - t_{28} + 4t_{24} &= -49.60/4 \\
4t_{40} + 4t_{36} + 3t_{32} + 2t_{28} + t_{24} &= 0
\end{aligned}$$

and

$$\begin{aligned}
r_0 + r_4 &= -3.81/16 \\
-r_0 - r_4 + 2r_8 &= -10.21/12 \\
-r_0 - r_4 - r_8 + 3r_{12} &= -16.28/8 \\
5r_0 + 4r_4 + 3r_8 + 2r_{12} &= 0.
\end{aligned}$$

The solutions to these equations are identical with those obtained from the normal least squares equations,

$$4(4t_{40} + r_0 + r_4 + r_8 + r_{12}) = T_{40}$$

etc.

The grand mean is added to the solutions to give the following values:

$$\begin{aligned}
t_{40} &= 7.48 & r_0 &= 6.31 \\
t_{36} &= 6.44 & r_4 &= 6.07 \\
t_{32} &= 5.59 & r_8 &= 5.77 \\
t_{28} &= 4.28 & r_{12} &= 5.37 \\
t_{24} &= 6.31
\end{aligned}$$

REFERENCES

Cochran, W. G., and Cox, G. M. *Experimental Designs.* John Wiley & Sons, 1950.

Yates, F. (1933). The principles of orthogonality and confounding in replicated experiments. *J. Agric. Sci., 23:* 108.

APPROPRIATE SCORES IN BIO-ASSAYS USING DEATH-TIMES AND SURVIVOR SYMPTOMS*

JOHANNES IPSEN
Associate Professor of Public Health, Harvard School of Public Health

Introduction

Death and survival are the most commonly used markers in biological standardization and in evaluation of medical therapy. Numerous methods are available for estimating the parameters of the dosage-mortality curve, and the biologist is often in a state of embarrassment of riches in the choice between a dozen well recommended methods to determine his LD50, ED50, TCiD50, or other "D50's" with which he is concerned.

The experimental records of good biologists and clinicians usually contain more information than death or survival of the subjects on given treatments, but either the "LD50-fixation" of the investigator prevents these data from entering in the evaluating process, or they are left out because no simple method is available to utilize these observations. In many assays, individual deaths have a quantitative connotation in terms of the time period from exposure to the lethal agent until death occurs. Survival has also quantitative aspects such as time of recovery, severity of symptoms at the time when death is no longer expected, etc.

Some data may consist mostly of deaths occurring at different times, in other experiments survivors are in excess. In these cases, group mortality percentages are too large or too small, respectively, to be of biometric use. Attempts are then made to find a transformation of death times or survivor symptoms that can be used as response metameter with approximately linear relationship to dose or with approximately normal distribution, or both. However, in such attempts the problem of truncated or censured distribution will sooner or later present itself, leaving the investigator either in uncomfortable indecision or involved in excessive computation.

*Presented at the Joint Meeting of the Institute of Mathematical Statistics and the Biometric Society, Chapel Hill, North Carolina, April 23, 1955.

This work was conducted under the sponsorship of the Commission on Immunization, Armed Forces Epidemiological Board, and was supported in part under contract with the Office of the Surgeon General, Department of the Army.

The most complex situation arises when some treatments of a series of graded treatments present groups with total mortality, and some present total survivorship, and intermediate treatments have partial mortality. The biologist, feeling that restriction to mortality percentages will throw some expensive observations out, may attempt to combine his graded observations on deaths and survivors in one continuous score system to include all data in the estimate of treatment effect. This writer has tried some graphical methods with the purpose of obtaining a normal distribution of such continuous score system in clinical (1) and laboratory data (2). In some immunological bio-assays a solution was attempted by assigning a score system (3), which—by graphical trial and error—gave a linear dose response curve. Both methods were unsatisfactory because there was no available method to show that the score systems utilized the biological data efficiently, and because the graphical methods were too subjective for reproduction by other workers.

Appropriate Scores for Linear Bio-Assays

The method to be presented provides a set of scores for multiple, mutually exclusive observations that satisfies *one* criterion for an efficient bio-assay:

The variance of the linear regression of the mean scores on log dose is the highest possible fraction of the total variance of a set of graded dose experiments.

Let us first consider an experiment where k different logarithmic doses x_1 , x_2 $\cdots$ x_k of the same preparation are given to k groups containing $a_{1.}$, $a_{2.}$ $\cdots$ $a_{k.}$ individuals. The observations are classes in m categories, e.g., ranging in death at 1, 2, 3 days and in severity of symptoms at survival. Subdivision in categories will be directed by biological experience and by the frequency of observation in each category.

If the number of individuals treated with log dose x_i that are observed in category j, is designated a_{ij} , the data may be arranged as follows (Table 1).

We shall use the sum of squares of deviations of the doses

$$S_{xx} = \sum_i x_i^2 a_{i.} - (\sum_i x_i a_{i.})^2 / a_{..} \tag{1}$$

and the normalized individual sums of products for each category

$$d_j = \sum_i x_i a_{ij} - \frac{a_{.j} \sum_i x_i a_{i.}}{a_{..}} \tag{2}$$

TABLE 1

log dose	Reaction category						Number Individuals per dose
	1	2	···	j	···	m	
x_1	a_{11}	a_{12}	···	a_{1j}	···	a_{1m}	$a_{1.}$
x_2	a_{21}	a_{22}	···	a_{2j}	···	a_{2m}	$a_{2.}$
·	·	·	···	·	···	·	·
·	·	·	···	·	···	·	·
x_i	a_{i1}	a_{i2}	···	a_{ij}	···	a_{im}	$a_{i.}$
·	·	·	···	·	···	·	·
·	·	·	···	·	···	·	·
x_k	a_{k1}	a_{k2}	···	a_{kj}	···	a_{km}	$a_{k.}$
Number per Category	$a_{.1}$	$a_{.2}$	···	$a_{.j}$	···	$a_{.m}$	Total $= a_{..}$

It is evident that

$$\sum_i d_i = 0 \tag{3}$$

If each category is given a score or measurement $z_1, z_2, \cdots z_j, \cdots, z_m$ belonging to a system (Z) we obtain the total sum of squares

$$S_{zz} = \sum_j a_{.j} z_j^2 - (\sum_j a_{.j} z_j)^2 / a_{..} \tag{4}$$

and the sum of products

$$S_{cz} = \sum_j (d_j z_j) \tag{5}$$

The problem is to find a system (C), $c_1, \cdots, c_j, \cdots, c_m$ which maximizes the function

$$\text{Max } \{r_{zz}^2\} = \text{Max} \left\{ \frac{(S_{cz})^2}{S_{zz} S_{zz}} \right\} = \frac{(S_{cz})^2}{S_{zz} S_{cc}} \tag{6}$$

It will be shown that such a system is obtained if

$$c_j = \frac{d_j}{a_{.j}} \tag{7}$$

or any linear transformation of c_j

Proof:

Two $(m-1)$ $(m-1)$ matrices are arranged as follows:

$$A = \begin{bmatrix} a_{.2} - \frac{a_{.2}^2}{a_{..}} & -\frac{a_{.3}a_{.2}}{a_{..}} & \cdots & -\frac{a_{.m}a_{.2}}{a_{..}} \\ -\frac{a_{.2}a_{.3}}{a_{..}} & a_{.3} - \frac{a_{.3}^2}{a_{..}} & \cdots & -\frac{a_{.m}a_{.3}}{a_{..}} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ -\frac{a_{.2}a_{.m}}{a_{..}} & -\frac{a_{.3}a_{.m}}{a_{..}} & \cdots & a_{.m} - \frac{a_{.m}^2}{a_{..}} \end{bmatrix}$$

$$D = \frac{1}{S_{xx}} \begin{bmatrix} d_2^2 & d_2d_3 & \cdots & d_md_2 \\ d_2d_3 & d_3^2 & \cdots & d_md_3 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ d_2d_m & d_3d_m & \cdots & d_m^2 \end{bmatrix}$$

A third matrix is constructed where each element is θ times the elements of the A-matrix minus the corresponding elements of the D-matrix. Equating the determinant of this matrix with zero, the following explicit solution is obtained:

$$\theta = \frac{1}{S_{xx}} \sum_i \frac{d_i^2}{a_{.i}} \tag{8}$$

We can now compute a score system (c') where $c_1' = 0$ and $c_m' = 1$ by solving $(m-2)$ sets of equations of the form

$$-c_2'\left(\frac{a_{.2}a_{.i}}{a_{..}}\theta + \frac{d_2d_i}{S_{xx}}\right) - \cdots + c_i'\left[\left(a_{.i} - \frac{a_{.i}^2}{a_{..}}\right)\theta - \frac{d_i^2}{S_{xx}}\right] - \cdots$$

$$- c_{m-1}'\left(\frac{a_{.i}a_{m-1}}{a_{..}}\theta + \frac{d_id_{m-1}}{S_{xx}}\right) = \frac{a_{.i}a_{.m}}{a_{..}}\theta + \frac{d_id_m}{S_{xx}}$$

The solutions are:

$$c_i' = \frac{1}{\frac{d_m}{a_{.m}} - \frac{d_1}{a_{.1}}}\left(\frac{d_i}{a_{.i}} - \frac{d_1}{a_{.1}}\right) \tag{9}$$

It is clear that the (C') system is linear with the (C) system as defined in (7) and hence *we can use* $c_i = d_i/a_{.i}$ *as the appropriate score system.*

Inserting (7) in (4) and (5) we obtain

$$S_{cx} = S_{cc} = \sum_i d_i^2/a_{.i} = \theta \cdot S_{xx} \tag{10}$$

and from (6)

$$\text{Max } \{r_{xz}^2\} = \frac{(S_{cx})^2}{S_{xx}S_{cc}} = \theta$$

Significance of θ

The square root of θ is the correlation coefficient between dose and scores and hence a significance test of the information extracted from a given experiment in respect to dose can be had from a table of r, entered with $(a_{..} - m)$ degrees of freedom.

Another test which is somewhat simpler, because it does not involve extracting a square root consists in assuming $a_{..}\,\theta$ to follow a χ^2-distribution with $m - 1$ degrees of freedom. This test is justified since $a_{..}\,\theta$ is the linear component in x of the χ^2 for an $(m) \times (k)$ contingency table.

We shall in the following use the term

$$\chi_c^2 = \theta a_{..} \tag{11}$$

Example 1. *Clinical scoring of typhus fever.*

Ecke et al. (4) found that persons immunized to varying degrees with typhus vaccine responded to subsequent typhus fever in various categories $A, B, \cdots, F$ classified in order of severity of clinical symptoms with F corresponding to death. The authors scored the six categories 0, 20, 40, 60, 80 and 100 with $A = 0$ and $F = 100$ and computed mean scores for each degree of immunization. These mean scores showed a good correlation with degree of immunization. (Table 2)

One may ask the questions: 1. What is the maximum information that can be obtained considering degree of immunization a linearly progressive scale? 2. What is the appropriate score system for the six categories under these conditions?

The cases were observed among the personnel in the typhus wards in Cairo during World War II. Immunization was attempted but some contracted the disease before completion of a three course dose schedule. The clinical categories were defined before the records were scrutinized.

Table 2 presents the data with "dose" ranked in equidistant classes $1 - 5$ (x). The computed appropriate scores that will form the best

TABLE 2

Example 1

Clinical severity of 71 cases of typhus fever related to degree of previous immunization (Ecke, et al. (4))

Degree of previous typhus immunization	Number of persons with typhus in clinical category (a_{ij})							Total	Mean Score	
	x_i	A	B	C	D	E	F	$a_{i.}$	Optimal c'	Assumed z
None	1			5	2		3	10	75.1	62.0
1 Dose < 12 days	2		4	8	4		1	17	57.0	43.5
1 Dose > 12 days	3		2	8	1			11	52.9	38.2
2 Doses	4		4	3				7	35.0	28.6
3 Doses	5	1	20	5				26	24.8	23.1
$a_{.j}$		1	30	29	7	0	4	71($a_{..}$)		
$\sum x_i a_{ij}$		1	130	82	13		5	235	$S_{xx} = \frac{11444}{71}$	
d_j (eq. 2)		$\frac{120}{71}$	$\frac{2180}{71}$	$\frac{-993}{71}$	$\frac{-722}{71}$		$\frac{-585}{71}$			
c_j (eq. 7)		1.69	1.02	−0.48	−1.45		−2.06			
$c'_j = \frac{c_j - c_A}{c_F - c_A} 100$		0	17.8	57.9	83.8		100			
z_j (assumed)		0	20	40	60		100			

(eq. 8) $\theta = 0.45148$ $r_{xz} = 0.672$ (D.F. = 66) $P < .001$

$a_{..}\theta = \chi^2_c = 32.055$ (D.F. = 4) $P < .001$

linear regression on x are shown in the row marked c_j. These scores are transformed linearly to c'_j so that direct comparison with the assumed scores z_j is facilitated.

The maximum information is highly significant whether the r-test or the χ^2-test is applied.

The appropriate scores are somewhat different from the assumed clinical score. As so often happens in medical science the death category was assigned a higher relative weight than the adjacent survivor categories.

A third and more interesting question is whether the assumed score

system was significantly less informative than the appropriate score system. This question needs a few theoretical considerations.

Test for "lost information".

Since the term $a_{..}\theta$ represents maximum information under the specified conditions, and

$$\chi_z^2 = a_{..}t_z = \frac{a_{..}}{S_{xx}}\frac{(S_{xz})^2}{S_{zz}} \tag{12}$$

represents information using the (Z) system, the difference

$$\chi_{c-z}^2 = a_{..}(\theta - t_z) = \frac{a_{..}}{S_{xx}}\left(\sum \frac{d_i^2}{a_{.i}} - \frac{S_{xz}}{S_{zz}}\sum d_i z_i\right) \tag{13}$$

measures the information which was lost due to use of the (Z) system.

Since $a_{..}\theta$ requires $(m-1)$ degrees of freedom and $a_{..}t_z$ only 1 degree of freedom, the lost information can be tested on the assumption that $a_{..}(\theta - t_z)$ is distributed as χ^2 with $m-2$ degrees of freedom.

In example 1, we have

	$a_{..}\theta = 32.055$	(D.F. = 4)	
	$a_{..}t_z = 28.082$	(D.F. = 1)	
"lost information"	3.973	(D.F. = 3)	$P = 0.28$

Hence, the assumed score system which was used by the authors did not cause a significant loss in information.

Assigning same score to adjacent categories.

The experimenter may start out with more observation categories than are biometrically significant. It may either be reflected in too small $a_{.i}$'s in some categories or in the fact that the computed score appears in a rank order that is biologically unsound — such as $c_j < c_{j+1} > c_{j+2}$ (where one would only accept $c_j < c_{j+1} < c_{j+2}$).

In these cases it is easy to form a combined score for adjacent categories.

$$c_{j,j+1,\cdots,j+p} = \frac{d_j + d_{j+1} + \cdots + d_{j+p}}{a_{.j} + a_{.j+1} + \cdots + a_{.j+p}} \tag{14}$$

A new ratio, θ', is computed with the following substitution:

$$\theta' = \frac{1}{S_{xx}}\left(\frac{d^2}{a_{.1}} + \cdots + \frac{d_{j-1}^2}{a_{.j-1}} + \frac{(d_j + \cdots + d_{j+p})^2}{a_{.j} + \cdots + a_{j+p}} + \cdots + \frac{d_m^2}{a_m}\right) \tag{15}$$

and the information lost by the combination of adjacent scores is esti-

mated by the difference

$$\chi^2 = a_{..}(\theta - \theta') \tag{16}$$

with degrees of freedom equal to the number of categories that were lost by the combination, in the above case equal to p.

Pooling scores from several experiments.

Computation of appropriate scores of a certain type of bio-assay is only useful if one can arrive at a set of scores that is appropriate for all experiments of the same type, and if it can be shown that one such set will cause no significant loss of information to any single experiment of the type.

If several experiments are carried out with the same type of reagents and observations are made under the same condition, a common score system for all experiments can be computed in the following way.

Observations are grouped in the same categories (j), and doses are expressed in logarithms; if coded, the code interval means the same log interval in all experiments.

Let a_{ijg} be the number of individuals observed for the ith dose, in the jth category in the gth experiment. Then

$$d_{i.} = \sum_g d_{ig}$$

$$\text{and } a_{.j.} = \sum_i \sum_g a_{ijg}$$

$$\text{The system } c_{j.} = \frac{d_{j.}}{a_{.j.}}$$

will then be the system of scores that maximizes the variance term of the common slope in relation to the total variance corrected for individual means. The test for the general application of this common score system will consist in applying the χ^2-test for lost information. In computing the several items for this test it is useful to conduct the following computational checks:

$$\sum_g S(c_{j.})a_{.jg} = 0$$

$$\sum_g S(c_{j.})^2 a_{.jg} = \sum_g Sd_{jg}(c_{j.})$$

By computing the common score system $(m - 2)$ degrees of freedom are used. Theoretically, this loss of d.f. would have to be distributed some way over all the experiments from which the score system is derived. If the number of experiments is large the individual χ^2-test will not be influenced by disregarding this difference.

Example 2.

Appropriate scores for bio-assay of tetanus toxoid in mice.

An inter-institutional study of the precision of tetanus toxoid assays in mice was reported a few years ago (5). This author who analyzed the data suggested a scoring system which was based on a graphical estimate that gave an apparent linear regression of mean score to log dose of toxoid. The set of 96 experiments made with various toxoids in six different laboratories provides an opportunity to apply the above method and test the validity of applying a uniform score system to all experiments.

The technique of the assay consisted in injecting graded doses of tetanus toxoid in groups of mice. Fourteen days later a measured dose of tetanus toxin was injected and deaths recorded daily for 7 days. On the seventh day survivors were recorded in two categories, survival with marked symptoms of tetanus, and survival without symptoms.

TABLE 3

Experiment I D-11. Tetanus toxoid assay.
Laboratory I. Toxoid *D*.

Dose of Toxoid	Coded log dose (x)	Number of mice dead on day						Survivors with tetanus	Survivors no sign	Total
		2	3	4	5	6	7			$a_{i.g}$
0.0125 ml	+1.5				1				5	6
0.00625 ml	+ .5	1	2					1	2	6
0.003125 ml	− .5	3	1					2		6
0.00156 ml	−1.5	4	1	1						6
	$a_{.ig}$	8	4	1	1	0	0	3	7	24 ($a_{..g}$)
	d_{ig}	−7.0	−1.0	−1.5	1.5	0	0	−0.5	8.5	0

$S_{xx} = 30.00$

Table 3 gives details of a protocol of one of the 6 experiments. There were five toxoids, one of which was assayed twice in each of 8 laboratory series. Each experiment was duplicated by random subdivision of each group of mice into two.

The designation of the experiments by ID11 means that it was part of laboratory I's series on toxoid *D*, first experimental day, first subgroup.

Table 3 also shows the value of the pertinent elements $a_{..}$, $a_{.i}$, S_{xx}

and d_j . These elements were summed over all 96 experiments and the resulting sums are shown in Table 4.

TABLE 4

Computation of common score system for 96 individual experiments

Category	No. mice ($a_{.j}$)	d_j.	Score	
			$c_{j.} = d_{j.}/a_{.j.}$	$(d_{j.})^2/a_{.j.}$
Death 2 days	500	−277.9407	−.5559	154.59
Death 3 days	211	− 64.7805	−.3070	19.89
Death 4 days	45	− 7.4286	−.1651	1.23
Death 5 days	28	+ 1.5000	+.0536	0.08
Death 6 days	29	+ 4.3787	+.1510	0.66
Death 7 days	26	− 9.4412	−.3631	3.43
Survivors with tetanus	427	+ 56.4701	+.1322	7.47
Survivors, no signs	524	+297.2422	+.5673	168.61
Sums	1790 ($a_{...}$)	0.0000		θ=355.87

$$S_{zz} = 1430.07$$

$$\chi^2 = a_{...}\theta = \frac{1790 \times 355.87}{1430.07} = 445.44 \text{ (D.F. = 7)}$$

The appropriate scores are shown in the fourth column ($c_{j.}$). Comparison with the observation categories shows that the score is increasing with prolonged survival except in the categories for deaths on 5, 6 and 7 days. Since this sequence is biologically unacceptable, and the numbers in these categories are small, it is logical to test whether the irregularity of the computed scores is merely the result of random variation.

Hence, a combined score is computed for a category comprising 5-7 days ($c_{5\text{-}7}$).

$$c_{5\text{-}7} = \frac{1.5 + 4.3787 - 9.4412}{28 + 29 + 26} = -\frac{3.5625}{83} = -.0429$$

This score follows the biological rank of the chosen categories. A test whether the combination of the three categories causes significant loss of information is indicated by equations (15) and (16) above.

$$\chi^2 = \frac{1790}{1430.07}\left(0.08 + 0.66 + 3.43 - \frac{(3.5625)^2}{83}\right) = \frac{1790 \times 4.02}{1430.07}$$

$$= 5.03 \qquad \text{(D.F. = 2)} \qquad (P = 0.08)$$

Thus, there is a probability of 0.08 that the irregularity of the three categories is due to random variation, and we can accept the new combined score for all three categories.

The score system (Z) which the author suggested in previous publications was as follows:

	assumed score (z_i)
Death in less than 2 days	0
Death in 3 to 4 days	2
Death in 5 to 7 days	3
Survival with tetanus	4
Survival with no signs	6

Testing the fit of this assumed system we compute,

$$\sum a_{.j.}(z_i) = 5613$$
$$\sum a_{.j.}(z_i)^2 = 27467$$
$$\sum d_{i.}(z_i) = 1854.2279 = S_{zz}$$

and

$$S_{zz} = 27467 - (5613)^2/1790 = 9866.01$$

$$a_{..}t_z = \frac{1790}{1430.07} \cdot \frac{(1854.2279)^2}{9866.01} = 436.19$$

$$a_{..}(\theta - t_z) = 445.44 - 436.19 = 9.25 \qquad \text{(D.F. = 6)}$$

The fit of the assumed score for all 96 experiments together is quite acceptable ($P = 0.16$). If the combined score for days 5-7 is used, the difference χ^2 is $9.25 - 5.03 = 4.22$ with 4 degrees of freedom. ($P = 0.39$). Since the earlier report indicated that the mean slope for each set of experiments varies significantly between laboratories, there are reasons to test whether the overall score systems ($C.$) and (Z) fit all laboratory sets without significant loss in information from that which would have been obtained by applying individual score systems (C_θ) for each laboratory.

Table 5 gives these individual systems transformed linearly so that scores for the highest and lowest category are 6 and 0, respectively. The variation in score from laboratory to laboratory for the four "free" categories is quite impressive, but the chi-square value of the last columns indicate that the variation from the overall score system ($C.$) indicated at the bottom of the table is not unreasonable. That of Laboratory V is the largest with $\chi^2 = 11.04$ ($P = 0.025$).

The assumed score system shows in 5 of 8 laboratories a poorer fit with two chi-square values (Lab. IVF and V) having probabilities less than 1%.

TABLE 5

Overall computed ($C_.$) and assumed (Z) score systems compared to systems (C_ϱ) computed for each of eight laboratory sets.

Laboratory	Mice used ($a_{..\varrho}$)	Score for category						χ^2 (C_ϱ) Indiv. System	Difference χ^2 for system		D.F.
		Death				Survival					
		2	3	4	5–7 days	symptoms	no sign		($C_.$)	(Z)	
I	240	0	2.07	1.40	2.68	3.21	6	86.38	1.24	0.99	4
II	215	0	2.82	6.22	5.40	6.18	6	22.43	4.36	3.39	4
III-1	216	0	1.38	−1.65	−0.90	2.23	6	46.70	5.52	6.91	4
III-2	214	0	1.51	2.90	−0.62	3.05	6	72.07	4.40	5.74	4
IV-*M*	213	0	0.64	−0.53	2.42	4.56	6	110.55	9.93	7.37	4
IV-*F*	216	0	0.26	1.46	1.00	2.08	6	87.26	6.37	15.02	4
V	263	0	0.62	5.37	6.18	3.98	6	94.54	11.04	13.77	4
VI	213	0	1.30	−0.07	4.18	4.37	6	34.03	2.61	2.70	4
Overall ($C_.$)	1790	0	1.33	2.09	2.74	3.68	6	440.41	—	—	
(Z)	1790	0	2	2	3	4	6	436.19	4.22	—	4

TABLE 6

Values of the Mean Slope for Each Laboratory as Computed with the Two Systems

Laboratory	Mean Slope*		Weight of Mean Slope	Intrinsic error		
	(C)	(Z)	S_{zz_g}**	$(C_.)$	(Z)	(C_g)
I	5.13	5.14	21.47	.403	.402	.399
II	2.79	2.81	12.86	.807	.785	.717
III-1	3.00	2.80	12.96	.505	.515	.466
III-2	5.57	5.50	12.86	.360	.366	.344
IV-*M*	6.42	6.42	12.77	.259	.253	.236
IV-*F*	5.77	5.53	12.96	.317	.346	.298
V	3.75	3.66	30.14	.496	.508	.452
VI	2.97	3.05	12.67	.586	.587	.559

*Increase in score per increase in log dose.
**x in log dose instead of coded values.

Table 6 gives the values of the mean slope for each laboratory as computed with the two systems. The difference is very slight. The last columns of Table 6 gives the effect on the precision of the assay caused by use of three score systems, as found in variation of the intrinsic error. This factor is usually computed as the square root of the remainder variance (s) over the slope (b). The values in the table have been computed with close approximation from the χ^2-values in the following way.

$$(s/b)^2 = \left(\frac{1}{\chi^2} - \frac{1}{a_{..g}}\right) S_{zz_g} \tag{17}$$

which is derived from the approximation

$$s^2 = \frac{1}{a_{..g}} \left(S_{yy} - \frac{S_{zy}^2}{S_{zz}}\right) \tag{18}$$

and the equations

$$b = \frac{S_{zy}}{S_{zz}} \tag{19}$$

and

$$\chi_s^2 = \frac{a_{..g} \cdot (S_{zy})^2}{S_{zz} \cdot S_{yy}}$$

Relative efficiency

It may be of interest to express the loss due to the assumed score-system in terms of efficiency. Since the relative efficiency of two methods of conducting a bio-assay is expressed by the reciprocal ratio of the values of $(s/b)^2$ obtained in each instance, we have a measure of relative efficiency in the term

$$E = \left(\frac{s_1}{b_1}\right)^2 \cdot \left(\frac{b_2}{s_2}\right)^2 \tag{20}$$

E will express the relative efficiency of method (2) as a fraction of that of method (1). Using this expression for two score systems (C) and (Z) we obtain

$$E = \frac{\chi_z^2(a_{..g} - \chi_c^2)}{\chi_c^2(a_{..g} - \chi_z^2)} = \frac{t_z(1 - \theta)}{\theta(1 - t_z)} \tag{21}$$

For example, if the overall system ($C_.$) is applied to the data for Laboratory I instead of the individual system (Cg) for that Laboratory, we find

$$E = \frac{85.14(240 - 86.38)}{86.38(240 - 85.14)} = 0.978$$

TABLE 7

Relative efficiency of assay when overall score systems ($C_.$) and (Z) are applied instead of individual score system (C_g).

Laboratory	Relative efficiency of	
	System ($C_.$)	System (Z)
I	0.978	0.982
II	0.788	0.834
III-1	0.854	0.819
III-2	0.911	0.885
IV-M	0.830	0.871
IV-F	0.883	0.741*
V	0.829	0.790*
VI	0.910	0.907

*Significant loss in information ($P < 0.01$).

Table 7 presents the relative efficiency for each laboratory when the computed overall system ($C_.$) and the assumed system (Z) is applied, instead of individual laboratory systems. In these experiments, apparently 20 per cent efficiency loss is not to be considered significant.

Distribution of χ^2_{c-s}

Since the assumed system (Z) can be considered adequate for most experiments it is of interest to examine the distribution of the statistic χ^2_{c-s} for each individual small experiment. Assuredly, the assumption that it follows the χ^2-distribution is being stretched rather far, considering that only 18 animals were used in most of the 96 subgroups experiments.

TABLE 8

Distribution of χ^2_{c-s} for 93 experiments.
Compared to the probability of the theoretical χ^2-distribution.

P	Degrees of freedom				Total	
	1	2	3	4	observed	expected
.0 – .01					0	0.9
.01– .05		1			1	3.7
.05– .10			2		2	4.7
.10– .20		3	5	2	10	9.3
.20– .50	4	10	15	5	34	27.9
.50– .80	1	15	13	3	32	27.9
.80– .90		2	5		7	9.3
.90– .95	2	1			3	4.7
.95– .99		2	2		4	3.7
.99–1.00					0	0.9
	7	34	42	10	93	93.0

Table 8 presents the distribution of χ^2_{c-s} for the number of degrees of freedom which exist considering the categories in which observation were found. In three experiments, only two categories contained observations; hence any system will fit these experiments and there are no degrees of freedom.

The distribution tends toward smaller values of χ^2_{c-s} than expected for a theoretical distribution, but considering the small number of animals on which the computations are based, the agreement is quite satisfactory.

The author wishes to acknowledge the able help of Mrs. Hanna Sylwestrowicz in performing the computations.

REFERENCES

1. J. Ipsen: Circulating antitoxin at the onset of diphtheria in 425 patients. *J. Immunology 54*, 325–347, 1946.
2. J. Ipsen and H.-I. Toft: Biological assay of strychnine in young mice using death time and convulsion time as biometrical scores. *Acta pharmacologica scand. 2*, 17, 1946.
3. J. Ipsen: The Effect of Environmental Temperature on the Immune Response of Mice to Tetanus Toxoid. *J. Immunology 69*, 273–283, 1952.
4. R. S. Ecke, A. G. Gilliam, J. C. Snyder, A. Yeomans, C. S. Zaraforetis, and E. S. Murray: The Effect of Cox-type Vaccine on Louse-borne Typhus Fever. *Amer. J. Trop. Med. 25*, 447–464, 1945.
5. J. Ipsen: Precision of Potency Assay of Alum-Precipitated Tetanus Toxoid in Mice. *J. Immunology 70*, 171–180, 1953.

MATRICES IN QUANTAL ANALYSIS

P. J. CLARINGBOLD
Department of Veterinary Physiology, University of Sydney, New South Wales, Australia

INTRODUCTION

Fisher (1954) has recently discussed the various transformations of probability used in the analysis of binomial data, and in that paper a full account of the statistical theory is given. While the assumption is often made that a distribution of thresholds must be postulated before efficient analysis may be made of binomial data (Finney, 1952a), Fisher (1954) has clearly demonstrated that this is unnecessary. Transformations of the expected proportion responding may thus be simply regarded as a different scale for the measurement of response. In this paper practical methods of relating the binomial variable to the coordinates of experimental designs are given, and matrix methods are employed so that the results may immediately be applied to any experiment with known design matrix.

A considerable time has been devoted in the past to methods supposed to give quick estimation of parameters in quantal analysis. These graphical or semigraphical methods are usually employed in order to avoid efficient but tedious probit analysis in routine work. In this paper it will be shown that quick efficient solution is afforded by use of the angular transformation.

Recently Berkson (1953) has advanced a "simplified and quick" method for the estimation of parameters of binomial data by means of a modified logit technique. The method is still tedious when compared with the method exemplified here, which gives estimates equivalent to those derived by Berkson's method.

NOTATION

Scalars—letters in italics.
Vectors—small letters in heavy type.
Matrices—large letters in heavy type.
Transposition of matrices or vectors is indicated by a prime.
Unit matrix—**1**.
Diagonal matrices—**diag** $(\cdots)$, indicating diagonal elements.

STATEMENT OF THE PROBLEM

An experimental design consists of N treatment combinations of k experimental variables or factors. If each of the factors is allocated a coordinate and scale of measurement the treatment combinations may be put in a one-one correspondence with row vectors of coordinate values, X_j , $j = 1, 2, \cdots, k$. Therefore N combinations may be represented in matrix notation by a *design matrix*, $\mathbf{D}$, where,

$$\mathbf{D} = [X_j^i], \qquad i = 1, 2, \cdots, N.$$

Successive rows of this matrix give the coordinate values of the N combinations.

These combinations of treatments are administered to N groups of animals, so that in each group some, all or none respond. The sizes of the N groups may be summarized in a column vector $\mathbf{n}$, the number responding in each group by the vector $\mathbf{a}$ and the observed proportion responding in each group in the vector $\mathbf{p}$. The ith element in each of these column vectors corresponds to the set of treatment combinations specified in the ith row of the design matrix. Likewise the expected proportions responding may be summarized in the vector $\boldsymbol{\pi}$, where $\boldsymbol{\pi} = E(\mathbf{p})$.

The statistical problem considered is the relation of the expected proportion responding to the coordinates of the experimental design. Suppose the scale of measurement of this proportion undergoes a continuous transformation into another scale. Thus the ith value of $\boldsymbol{\pi}$ is a function of the ith value of some transformate, i.e.,

$$\pi_i = f(\rho_i). \tag{1}$$

The vector of N values of ρ_i is denoted $\boldsymbol{\rho}$. The Jacobian matrix of this transformation $\partial\boldsymbol{\pi}/\partial\boldsymbol{\rho}$, is diagonal and denoted $\mathbf{diag}.(J)$.

The transformed vector may be related to the coordinates of the experimental design in a manner very similar to the standard regression analysis adopted with the normally distributed variable. The maximum likelihood procedure is outlined in the Appendix and is preferred since statistics estimated by this means have asymptotically normal distributions and the standard tests of significance may therefore be applied with some confidence. At the beginning of a regression analysis it is decided to relate expected response to certain functions of the coordinates of the experimental design. These may be powers, cross-products, certain sets of orthogonal functions and other functions. It is assumed that response may be expressed as a linear combination of these and a set of regression coefficients, $g' \leq N$ in number, $(g' = g + 1)$.

The coordinate functions may be evaluated at each point of the experimental design, that is for each row of the design matrix. The N by g' set of values may therefore be arranged as a matrix called the *matrix of coordinate functions.* It is assumed that the rank of this matrix is g', or that, in other words, the coordinate functions are linearly independent. Thus the ith value of the response transformate is a linear combination of a set of regression coefficients and the elements of the ith row of the matrix of coordinate functions.

$$\rho^i = x_0^i\beta_0 + x_1^i\beta_1 + \cdots + x_g^i\beta_g \tag{2}$$

or $\boldsymbol{\rho} = \mathbf{X}\boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is a vector of regression coefficients. Since this transformation is linear it has the matrix $\mathbf{X}$ as its Jacobian, i.e. $\partial\rho^i/\partial\beta_s = x_s^i$, where x_s^i is the i, sth element of $\mathbf{X}$, the matrix of coordinate functions. In equation (2) small letters x are used to distinguish the values of the coodinate functions from those of the original coordinates X_i. In general the sth coordinate function x_s may be a function of any, some or all of the coordinates. One function is always defined, $x_0 = 1$, and is termed the identity. Estimates of β_0 give the response intercept when all other coordinate functions are zero.

Provided that it is possible to describe the response in terms of the functions chosen the solution is readily obtained in practice. Otherwise the iterative procedure used may take many steps and the regression coefficients estimated are biassed. This feature of regression analysis is fully discussed by Box and Wilson (1951) where the General Theory of Aliases is given in matrix notation.

EXAMPLES

Examples involving a small number of experimental points must be chosen if they are to go on a page of the journal. The method of analysis and the laying out of computations in larger examples are identical in form to those illustrated, simply occupying more space. In the practical carrying out of certain matrix multiplications the computor needs to learn one new operation, namely, row into row or column into column multiplication.

Examples with diagonal information matrix.

Into this class fall all factorial and other orthogonal designs where the angular transformation is employed with equal group sizes. In certain other designs the information is readily reducible to the diagonal form by means of suitable coordinate transformations, for example the parallelogram designs of Claringbold, Biggers and Emmens (1953)

1. *Four point assay.*

Berkson (1953) advocates a "short cut minimum logit χ^2 method" for the four point assay. The analysis takes a full page of the journal in small type and requires reference to tables and nomographs not freely available. In Berkson's example the ratio of successive doses of the standard preparation (S) and the unknown preparation (U) was 1.5.

$1/nw$ $(\mathbf{X}'\mathbf{X})^{-1}$ $\mathbf{X}' = [x_i']$

					S		U		$\check{\beta}$	$\beta_{(1)}$	s_β
820.7/24	0.25	·	·		1	1	1	1	β_0 47.0	46.83	2.93
	·	0.25	·		−1	−1	1	1	β_1 4.5	4.33	2.93
	·	·	0.25		−1	1	−1	1	β_2 15.5	15.43	2.93

	$\mathbf{p}'$	25	67	29	88	Log M = −0.2906
$\chi^2_{(1)} = 1.06$,	$\mathbf{r}'$	30	55	33	70	± 0.1973
$P > 0.05$	$\mathbf{e}'_{(1)}$	27	58	36	67	M = 0.893
	$\psi'_{(1)}$	30.1	54.9	32.7	69.6	· {0.760 − 1.039}

The design matrix may be seen in the second and third rows of $\mathbf{X}'$. The regression coefficient β_0 estimates the mean response, β_1 the difference between samples and β_2 the slope of the dose response lines. The computational operations comprise:

(i) Form $\mathbf{X}'$: At most four regression coefficients are required to describe the experimental data. Three have been estimated above while the fourth could be a measure of the departure from parallelism between the two dose response lines.

(ii) Form $(\mathbf{X}'\mathbf{X})^{-1}$: Since the coordinate functions are orthogonal this matrix is diagonal. The orthogonality of the functions may be checked by seeing that the sum of the cross products of equivalent elements in different rows is zero. The diagonal elements of $(\mathbf{X}'\mathbf{X})^{-1}$ are the reciprocals of the sums of squares of the individual rows of $\mathbf{X}'$. The scalar $1/nw$ may immediately be tabulated, and together with $(\mathbf{X}'\mathbf{X})^{-1}$ gives the variance-covariance matrix of the regression coefficients. Thus the standard errors of these coefficients may be immediately determined and tabulated. For example,

$$s(\beta_0) = \left\{\frac{820.7}{24} \times 0.25\right\}^{1/2}$$

(iii) Write the observed proportions responding under the appropriate column of $\mathbf{X}'$, i.e. complete the row labelled $\mathbf{p}'$.

(iv) Use Fisher and Yates (1953) Table XII or Finney (1952b) Appendix Table XI to tabulate the empirical angles as a row $\mathbf{r}'$.

(v) Calculate provisional estimates of the regression coefficients by row into row multiplication, thus

$$x_1^s r_1 + x_2^s r_2 + x_3^s r_3 + x_4^s r_4 , \qquad s = 0, 1, 2.$$

and multiply by the sth element of $(\mathbf{X}'\mathbf{X})^{-1}$ to give β^s. For example $47.0 = (1 \times 30 + 1 \times 55 + 1 \times 33 + 1 \times 70) \times 0.25$

(vi) Form the row of expected responses by column into column multiplication, thus

$$\rho^i = x_i^1\beta^1 + x_i^2\beta^2 + x_i^3\beta^3 + x_i^4\beta^4, \qquad i = 1, 2, 3 \text{ or } 4.$$

For example, $27 = (1 \times 47.0) - (1 \times 4.5) - (1 \times 15.5)$

(vii) Use Finney (1952b) Appendix Table XIII to tabulate, or Fisher and Yates (1953) Table XIV to calculate the row of working angles, ψ'.

(viii) Repeat operation (v) using the working angles. The iterative procedure may be considered complete if the 10% criterion is adopted (see below). Otherwise operations (vi) and (vii) are alternated, building up new rows of working angles and new columns of estimates.

(ix) Log relative potency, $\log M$, is given by the ratio β_1/β_2 , and a negative sign is taken since the response to U is less than to S.

(x) The variance of $\log M$, $V(\log M)$, is given by Fieller's formula, discussed by Finney (1952b). Since β_1 and β_2 are independent, in this example, the formula simplifies,

$$V(\log M) = 1/\beta_2^2 \cdot [V(\beta_1) + (\log M)^2 V(\beta_2)]$$

$$= 820.7/(96\beta_2^2) \cdot [1 + (\log M)^2]$$

(xi) Fiducial limits are obtained in the log scale and then transformed into the arithmetic scale by taking antilogarithms to the base 1.5.

(xii) The χ^2 goodness-of-fit is formed from the difference of the weighted sum of squares of the working angles and the weighted sum of cross

products between the working angles and the expected response. It has $4 - 3 = 1$ degree of freedom. Thus

$$\chi^2 = [(30.1^2 + 54.9^2 + \cdots) - (30.1 \times 27.1 + 54.9 \times 57.9 + \cdots)] \cdot 24/820.7 = 1.06$$

The calculation of relative potency is unnecessary in this example if all that is required is the answer to the question, "Does the activity of U differ significantly from that of S?" The significance of β_2 tests this hypothesis,

$$\text{i.e.} \qquad t_{(\infty)} = 1.48, \qquad P > 0.05$$

Berkson's estimate of M for this example is 0.887 {0.746 − 1.056}, which differs by only 0.006 from ours. This is a negligible difference in terms of the confidence interval. Berkson (1953) mistakenly uses the antilogarithm of the standard error of log relative potency as standard error of relative potency. The fiducial limits are the appropriate measure of accuracy here. It may be noted that the angular transformation gave a narrower fiducial interval than logits.

2. *Parabola design.*

Examples of parallelogram designs have been given by Claringbold, Biggers and Emmens (1953). In that paper it was suggested that the principle developed could be extended. The present example is extracted from a 2×5^2 factorial, for illustrative purposes, and has the design matrix contracted into groups of three rows:—

$$\mathbf{D} = \begin{array}{c} \begin{matrix} X_1 & \overbrace{\qquad X_2 \qquad} \end{matrix} \\ \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 2 \end{bmatrix} \end{array}$$

Thus at the zero level of X_1, the levels of $\log_2$ dose, namely X_2, were lowered by one unit. The design presupposes a quadratic relationship between response and X_1, it being known in advance that the middle level of X_1 would markedly increase response to dose.

It was decided to relate response to four coordinate functions, $x_0 = 1$, $x_1 = X_1$, $x_2 = (X_1)^2$ and $x_3 = X_2$. The transposed matrix of coordinate functions is,

$$\mathbf{X}'$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 0 & 1 & 0 & 1 & 2 \end{bmatrix} \begin{matrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{matrix}.$$

Formation of the sums of squares and cross products between pairs of rows of this matrix show these functions to be non-orthogonal. The variance covariance matrix is therefore non-diagonal.

Premultiplication of the matrix $\mathbf{X}'$ by the matrix $\mathbf{C}'$ where,

$$\mathbf{C}' = \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ -2 & \cdot & 3 & \cdot \\ \cdot & \cdot & -1 & 1 \end{bmatrix}$$

gives the transformed matrix of coordinate functions, $\mathbf{X}^{*\prime}$, tabulated below. The computational procedures are similar to those given above for the four point assay except for the stages after iteration when the unstarred coordinate functions are reintroduced.

In this analysis the 100% values are initially given the value 84 in conformity with the correction discussed by Claringbold, Biggers and Emmens (1953) and used with the empirical angular transformation. The χ^2 goodness of fit is not significantly enlarged and it may therefore be assumed that the regression model adequately described the data, i.e.

$$\rho = 46.1 - 2.9x_1^* + 3.0x_2^* + 32.1x_3^*$$

The vector of estimated regression coefficients may be transformed so that they relate response to the original unstarred coordinate functions, thus:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 & \cdot & -2 & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 3 & -1 \\ \cdot & \cdot & \cdot & 1 \end{bmatrix} \cdot \begin{bmatrix} 46.1 \\ -2.9 \\ 3.0 \\ 32.1 \end{bmatrix} = \begin{bmatrix} 40.1 \\ -2.9 \\ -23.1 \\ 32.1 \end{bmatrix}$$

or $\beta = C\beta^*$, see Appendix equation 8.

$1/nw$	$(\mathbf{X}^{*\prime}\mathbf{X}^{*})^{-1}$				$\mathbf{X}^{*\prime}$							$\hat{\beta}^{*}$	$\hat{\beta}^{*}_{(1)}$	$\hat{\beta}^{*}_{(2)}$	$\hat{\beta}^{*}_{(3)}$	s^{2}_{β}	
820.7/24	0.1111	.	.	.	1	1	1	1	1	1	1	45.3	46.1	46.3	46.1	2.0	x_0^{*}
	.	0.1667	.	.	−1	−1	−1	0	0	0	1	−2.3	−2.7	−2.9	−2.9	2.4	x_1^{*}
	.	.	0.0556	.	1	1	1	−2	−2	−2	1	3.3	3.1	2.9	3.0	1.4	x_2^{*}
	.	.	.	0.1667	−1	0	1	−1	0	1	−1	32.3	32.2	31.9	32.1	2.4	x_3^{*}

p'	8	63	100	8	25	88	4	46	100	
r'	16	53	84	16	30	70	12	43	84	
$\rho'_{(1)}$	19	51	83	6	39	71	14	46	79	
$\psi'_{(1)}$	16.6	52.5	86.5	19.5	30.5	69.7	11.8	42.8	84.6	
$\rho'_{(2)}$	19.7	51.9	84.1	7.7	39.9	72.1	14.3	46.5	78.7	
$\psi'_{(2)}$	16.7	52.6	87.1	21.1	30.6	69.6	11.8	42.7	84.4	
$\rho'_{(3)}$	20.2	52.1	84.0	8.6	40.5	72.4	14.4	46.3	78.2	
$\psi'_{(3)}$	16.7	52.6	87.0	19.8	30.6	69.6	11.8	42.8	84.2	

$$\chi^{2}_{[6]} = 8.1, \quad 0.3 > P > 0.2$$

Thus

$$p = 40.1 - 2.9x_1 - 23.1x_2 + 32.1x_3$$
$$= 40.1 - 2.9X_1 - 23.1(X_1)^2 + 32.1X_2$$

Since the starred regression coefficients are independent the variances of any linear combinations of them are simply determined. For example since $\beta_0 = \beta_0^* - 2\beta_2^*$, then $V(\beta_0) = V(\beta_0^*) + 4V(\beta_2^*)$.

Example with non-diagonal information matrix.

The type of analysis demonstrated below must be used with all non-orthogonal and/or badly designed experiments or when group sizes are unequal. The process must also be used with all transformations other than angular with the added disadvantage that the variance-covariance matrix changes, albeit slowly, at each stage of iteration. The example chosen is a 2^3 factorial experiment with unequal group sizes. One factor is replication so that the parameters requiring estimation are the identity, replicate difference, the effect of each treatment and the interaction between treatments i.e. five regression coefficients. The coordinate functions are $x_0 = 1$, $x_R = X_R$, x_1, $= X_1$, $x_2 = X_2$, $x_3 = X_1X_2$.

The computational procedures comprise:

(i) Determine a working unit of weight. In this example the average group size was given unit weight.

(ii) Write down the matrix coordinate functions, **X′**.

(iii) Multiply successive columns of this matrix by the appropriate group weight. This is tabulated unchanged in the first row of the weighted matrix of coordinate functions, **X′W**.

(iv) Form the matrix **X′WX** by row into row multiplication of **X′** into **X′W**. Thus the second row, third column element of **X′WX** is obtained by the row 2 into row 3 multiplication, viz.

$$(-1 \times -1.08) + (-1 \times -1.08) + (-1 \times 0.92) + (-1 \times 1.03)$$
$$+ (1 \times -1.08) + (1 \times -1.08) + (1 \times 0.70) + (1 \times 1.03) = -0.22$$

(v) Form the inverse of this matrix **X′WX** using the method of say Fox (1950) and Fox and Hayes (1951). This involves formation of an upper triangular matrix (**A**) and then the inverse. The full layout is shown, excluding check lines, but will not be explained as this is fully carried out by Fox and Hayes (1951). The inverse matrix is the

Design matrix		Ratio no. responding to no. per group	
Factor A	Factor B	First Replicate $X_R = -1$	Second Replicate $X_R = 1$
X_1	X_2		
−1	−1	3:*20*	4:*20*
−1	1	15:*20*	15:*20*
1	−1	4:*17*	5:*18*
1	1	17:*19*	17:*19*

Average group size = 18.50

$\mathbf{X}'$ | $\mathbf{X}'\mathbf{W}$

x_0	1	1	1	1	1	1	1	1	1.08	1.08	0.92	1.03	1.08	1.08	0.70	1.03
x_R	−1	−1	−1	−1	1	1	1	1	−1.08	−1.08	−0.92	−1.03	1.08	1.08	0.70	1.03
x_1	−1	−1	1	1	−1	−1	1	1	−1.08	−1.08	0.92	1.03	−1.08	−1.08	0.70	1.03
x_2	−1	1	−1	1	−1	1	−1	1	−1.08	1.08	−0.92	1.03	−1.08	1.08	−0.70	1.03
x_3	1	−1	−1	1	1	−1	−1	1	1.08	−1.08	−0.92	1.03	1.08	−1.08	−0.70	1.03

X'WX

8.00	−0.22	−0.65	0.43	0.43
−0.22	8.00	−0.22	0.22	0.22
−0.65	−0.22	8.00	0.43	0.43
0.43	0.22	0.43	8.00	−0.65
0.43	0.22	0.43	−0.65	8.00

A

2.83	−0.08	−0.32	0.15	0.15
0	2.83	−0.08	0.08	0.08
0	0	2.82	0.17	0.17
0	0	0	2.82	−0.25
0	0	0	0	2.81

1000(X'WX)$^{-1}$

127	4	12	−8	−8
4	126	4	−4	−4
12	4	127	−8	−8
−8	−4	−8	127	12
−8	−4	−8	12	127

1000T

120	120	138	121	130	130	112	129
−136	−136	−100	−129	136	136	100	129
−129	−130	138	121	−120	−121	112	129
−118	128	−141	131	−128	118	−113	123
127	−118	−140	131	118	−128	−113	123

$\mathbf{p}'$	15	75	24	90	20	75	38	90
$\boldsymbol{\gamma}'$	23	60	29	72	27	60	38	72
$\boldsymbol{\varrho}'_{(1)}$	24	58	31	71	27	61	34	72
$\boldsymbol{\psi}'_{(1)}$	22.8	60.0	29.4	71.6	26.6	60.0	38.2	71.4
$\boldsymbol{\psi}'_{(1)}\mathbf{w}$	24.6	64.9	27.0	73.5	28.8	64.9	26.9	73.3

$\check{\boldsymbol{\beta}}_w$	$\boldsymbol{\beta}_{(1)}$	$\beta_{(2)}$	S_β
47.5	47.4	2.4	x_0
1.5	1.4	2.4	x_R
5.0	5.0	2.4	x_1
18.5	18.4	2.4	x_2
1.2	1.1	2.4	x_3

$x^2_{(3)} = 2.11, \quad 0.7 > P > 0.5$

variance covariance matrix when multiplied by a scalar 820.7 ÷ 18.5. i.e. the reciprocal of the unit weight.

(vi) Form the transformation matrix (**T**) by column into column multiplication of the inverse matrix (excluding the scalar factor since this cancels out) into the weighted matrix of coordinate functions, **X'W**. Thus the first row, second column element of **T** is given by:

$$(0.127 \times 1.08) + (0.004 \times -1.08) + (0.012 \times -1.08) + (-0.008 \times 1.08) + (-0.008 \times -1.08) = 0.120.$$

In all the above matrix operations two additional decimal places were held at all stages. The figures have been rounded to save space. The sum of the elements of the first row of **T** should be unity and the remainder zero.

(vii) Write down row of observed percentage responding under each appropriate column of **T**.

(viii) Begin the iterative procedure in an exactly analogous manner to the above examples, obtain regression coefficients by row into row multiplication of **T** into the working of empirical angles, and new rows of expectation by column into column multiplication of regression coefficients into the matrix of coordinate functions.

(ix) The calculation of χ^2 goodness-of-fit requires an additional step at the end of iteration. A row is formed the elements of which are the product of the final working angles and the corresponding weight (first line of **X'W**). The sum of cross products of the elements of this and the previous line, minus the sum of cross products between the elements of this line and the elements of the line of expected responses above, multiplied by the value of the unit weight, gives a χ^2 goodness-of-fit test. Thus

$$\chi^2 = (24.6 \times 22.8 + \cdots) - (24.6 \times 24 + \cdots) \times \frac{18.5}{820.7}$$

$$= 2.11$$

DISCUSSION

The angular transformation may be used in two distinct types of experimental problem where the response variable is binomial. The first type of problem is exemplified by studies of some dose response law. If it is of interest to show a log normal tolerance distribution studies must be made with the probit transformation. If such a law has been established then the angular transformation may only be used as a very good approximation in the interval of 5–95% expected response. This has been practically demonstrated by Biggers (1951) and may be

observed by graphical comparison of the transformations (Finney, 1952b). Similar considerations apply to logits.

In the second type of problem such a response law may be unknown or not the prime interest of study. For example Cochran (1938) discusses field experiments with percentage data and does not mention the tolerance distribution. More recently Campbell, Hancock and Rothschild (1953) have applied the angular transformation to percentage dead or alive spermatozoa using the angular transformation simply as a convenient scale. Finally Fisher (1954) does not mention tolerance distributions but regards transformations simply as transformations of probability of response. In this type of problem the angular transformation is not an approximation but is a more convenient measure of response.

Various binomial transformations are used in the case where variation is quantal but in excess of the Bernoulli binomial distribution. The problem has been recently discussed by Bartlett (1954) and Anscombe (1954). In this case the distribution is unknown and the maximum likelihood procedure is inapplicable. While the angular transformation may be used here any procedures are largely arbitrary. Bartlett (1954) states that there are very few examples of binomial data in the literature which are not heterogeneous. In this laboratory it has been found that with strict randomisation of animals to experimental units and with stratified randomisation of operators to the experimental units the conditions for obtaining homogeneous results occur. At present thirty factorial or other complex experiments have been carried out with the binomial variable, analysed with the angular transformation and all but one found homogeneous. These are available on request, and have been published in a number of biological journals (Biggers & Claringbold, 1954; Claringbold, 1953). It seems that strict randomisation and control is essential.

The first two examples illustrate the simplicity of the estimation procedure advocated in this paper. Berkson (1953) makes an important point when he criticises the routine users of probit analysis who employ only one cycle of the iterative procedure following graphical estimation. The procedure leads to widely discrepant results in small samples, say of 50 animals. The time taken in routine analysis with the probit transformation is excessive and a frequent cause of this error, while in this type of work use of the angular transformation with equal steps in the log dose scale results in quick analysis. Groups of animals with expected responses outside the interval 5–95% may be ignored to preserve linearity.

When should iteration be stopped? One arbitrary but stringent

criterion suggested by Fisher and Yates (1953) is the acceptance of estimates which differ from those of the previous iteration, or the previous graphical estimates in the case of the first iteration, by less than 10% of their standard error. It is difficult to imagine any experimental situation where differences of this order could be considered important, after all, a statistic with infinite degrees of freedom must exceed zero by 196% of its standard error before being considered significant.

Large regions of 0 and 100% responses in factorial experiments have been discussed by Claringbold, Biggers and Emmens (1953). It was concluded that the best method was to use special experimental designs based on small scale pilot experiments. The second example is a simplified form of such a special design. The aim of the design was to obtain a precise estimate of slope with three levels of an additional variable. The doses were therefore spaced as widely as possible in the interval 5–95%. Another aim of this design was to establish more precisely the quadratic relationship of response to X_1 .

Alternatives to this design have disadvantages. The experiment could be designed as a 3^2 factorial with the dose interval halved so the responses avoided the extremes of the scale. This results in a loss of 75% of the information about slope. If a standard factorial were used without the reduction in scale interval there would be, depending on the centring, two groups with expectations very near 100% on the middle line or two groups with expectations near 0% on the outer lines. Linearity in angles would therefore be lost and to preserve linearity, analysis in terms of probits or logits would have to be made with resultant loss in simplicity.

In probit and logit analysis of factorial or other experiments where there are large regions of 0 and 100% response an additional difficulty arises. When the expected response approaches 0 and 100% the weight of a probit or logit approaches zero, and may reach zero in terms of the number of figures carried by the computer. If the number of experimental groups giving reasonable amounts of information is less than the number of regression coefficients requiring estimation, attempts to fit this number of regression coefficients are doomed to inaccuracy and near aliases. This is evinced by the misbehaviour of the numerical computation when logits or probits are used in such cases. The mathematical reason underlying this situation is that the rank of the information matrix becomes less than g', and before it may be inverted the linearly related coordinate functions must be eliminated. It can be seen therefore that problems of regions of 0 and 100% responses are not

confined to the angular analysis and are best avoided by designing the large scale experiment on the basis of small pilot tests.

When the expectation of every group is at one end of the response scale the number of animals responding (or alternatively, those not responding, whichever is the smaller may be regarded as a Poissonian variable and the square root transformation used. This transformation has the same effect as the angular in equalising information.

With the third example the awkwardness of non-orthogonal or unequally weighted data is illustrated. This whole process must theoretically be repeated at each cycle for transformations other than the angular, since the information changes with each cycle. Usually it is sufficient to calculate information initially from the trial solution, run through the iterative procedure to convergence and then recompute information and check convergence.

While the first example took 11 minutes and the second 33 minutes to analyse, including checks, the third involved some 126 minutes of computational time, most of which was spent in forming and inverting the information matrix. If this had been repeated for different transformations the information matrix would need to be formed at least twice thus involving a tedious 4 or more hours work. This is a small factorial experiment. Consider the analysis of a 2^6 experiment in say probits. If we wish to estimate a mean, all main effects and first order interactions, 22 parameters require estimation. It would appear that the inversion of the 22 by 22 information matrix renders probit analysis prohibitive.

Finney (1952a) in the discussion of a 2^3 factorial experiment used by Potter and Gillham (1946) suggests a different approach. In this experiment a dose response line was determined using five doses under the eight sets of conditions specified in the 2^3 design. The method of analysis was: (1) compute the eight dose response lines using probits, (2) test for parallelism of the lines and their homogeneity, and estimate a common slope, (3) compute the weighted means ($\bar{x}$ and $\bar{y}$) of each dose response line, (4) perform an *unweighted analysis* of these means to obtain estimates of the eight parameters required to describe each of these, i.e. grand mean, three main effects, three first order interactions and a second order interaction. All but the first of these parameters is a difference between four weighted means and the other four. (5) Divide these seven differences of the $\bar{y}$ by the common slope and subtract from the corresponding difference of the $\bar{x}$, to give relative potency figures to which may be attached a standard error. This analysis is still tedious since eight lines must be fitted and iterated one at a time.

The most serious objection is the unweighted analysis of the weighted means. Their weights ranged from 24.2 to 115 and so the efficient analysis must be a weighted analysis. The approach runs into further complications when slope is affected by the treatments (see for e.g. Biggers, Claringbold and Emmens, 1954), when a separate weighted analysis of slope must be carried out. Unfortunately this example is heterogeneous and many 100% responses are present, thus rendering it inappropriate for angular analysis.

It may be concluded that only with the angular transformation are the modern experimental designs to be applied freely with the binomial variable. It cannot however, be concluded that this transformation will be the best or appropriate in all experimental situations; it may only be said that it appears by far the simplest solution in laboratory experiments where rigid control and randomisation ensure the classical Bernoulli binomial distribution. Some have said that use of the angular transformation may give misleading results. The final arbiter on this point is the goodness-of-fit test. If this is satisfactory our assumptions of both transformation and form of the regression equation have not been disproved.

APPENDIX

Fisher (1954) has given a full account of the binomial probability distribution, most transformations commonly used, the theory of the maximum likelihood procedure and its application to the present problem. The matrix equivalent equations are developed briefly here in order to justify the estimation procedure laid out in the examples.

Likelihood equations and information

The loglikelihood (Fisher, 1925, 1934) of the vector $\mathbf{a}$ defined above is given by,

$$L = \mathbf{a}'[\log \pi_i] + (\mathbf{n}' - \mathbf{a}')[\log (1 - \pi_i)] \tag{3}$$

The quantities in square brackets are column vectors of length N, the ith element of which is shown.

Thus loglikelihood is a function of the observations $\mathbf{a}$ and the expected proportion $\boldsymbol{\pi}$ apart from a constant which has been ignored. The expected proportion is a function of the transform ($\boldsymbol{\rho}$), which in turn is a function of the vector of regression coefficients, $\boldsymbol{\beta}$.

The maximum likelihood estimate is obtained by the solution of the partial differential equations, $\partial L/\partial \boldsymbol{\beta} = 0$. In terms of the transformations (1) and (2), as a column vector,

$$\partial L/\partial \boldsymbol{\beta} = \mathbf{X}' \operatorname{diag}(J)\, \partial L/\partial \boldsymbol{\pi} = 0 \tag{4}$$

The information in the sample with respect to the parameters is given by the negative of the expectation of the second partial differentials of the loglikelihood equation with respect to the parameters.

Thus $\mathbf{I} = \mathbf{X'WX}$ where $\mathbf{W}$ is the diagonal weight matrix of the transformation, defined,

$$\mathbf{W} = \{\mathbf{diag}(J)\}^2 \cdot \mathbf{diag}\{n_i/\pi_i(1 - \pi_i)\} \tag{5}$$

The variance covariance matrix of the regression coefficients is given by,

$$\mathbf{V} = \mathbf{I}^{-1}, \qquad \text{provided } [\mathbf{I}] \neq 0.$$

Solution of the estimation equations

The solution is based on the scoring system discussed by Fisher (1946) and described in the general case by Rao (1952). The vector of first partial differentials $\partial L/\partial \beta$ is expanded about a trial value indicated $\partial L/\partial \beta_{(0)}$ to the first order. In this equation the information matrix may replace the second partial differentials so that,

$$\partial L/\partial \beta = \partial L/\partial \beta_{(0)} + \mathbf{I}_{(0)} \Delta \beta_{(0)} = 0 \tag{6}$$

where $\Delta \beta$ is a vector of additive corrections. The bracketed subscripts denote evaluation at a trial value.

Solution is facilitated by the introduction of a *working vector variate* which may be defined in a number of alternative ways. For example,

$$\psi = \{\rho - \mathbf{diag}(J)^{-1}\pi\} + \mathbf{diag}(J)^{-1}\mathbf{p},$$

where the quantity in braces is called the *minimum working variate* and the individual values of $\mathbf{diag}(J)^{-1}$ the *range*.

Substitution in equation (6) with considerable rearrangements gives, at the ith stage of iteration,

$$\begin{aligned} \beta_{(i+1)} = \beta_{(i)} + \Delta\beta_{(i)} &= (\mathbf{X'W}_{(i)}\mathbf{X})^{-1}\mathbf{X'W}_{(i)}\psi_{(i)} \\ &= \mathbf{T}_{(i)}\psi_{(i)} , \end{aligned} \tag{7}$$

where $\mathbf{T}$ is termed the *transformation matrix*.

Linear transformations of the matrix of coordinate functions.

Suppose in equation (2) a linear transformation of rank g' is made on the matrix of coordinate functions by the square matrix $\mathbf{C}$. The inverse transformation is made on the vector of regression coefficients, i.e.

$$\rho = \mathbf{XCC}^{-1}\beta = \mathbf{X}^*\beta^* \tag{8}$$

The maximum likelihood estimate of the transformed vector may be reached in the standard manner as outlined above. On reaching the solution the inverse transformation may simply be made from the relation $\boldsymbol{\beta} = \mathbf{C}\boldsymbol{\beta}^*$. Thus the matrix $\mathbf{C}^{-1}$ need never be computed. This procedure will not affect estimates of regression coefficients as is shown,

$$\begin{aligned}\boldsymbol{\beta} = \mathbf{C}\boldsymbol{\beta}^* &= \mathbf{C}(\mathbf{X}^{*\prime}\mathbf{W}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\mathbf{W}\boldsymbol{\psi} \\ &= \mathbf{C}(\mathbf{C}'\mathbf{X}'\mathbf{W}\mathbf{X}\mathbf{C})^{-1}\mathbf{C}'\mathbf{X}'\mathbf{W}\boldsymbol{\psi} = \mathbf{C}\mathbf{C}^{-1}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{C}'^{-1}\mathbf{C}'\boldsymbol{\psi} \\ &= \text{equation 7.}\end{aligned}$$

While it is always possible to find such a transformation which will render the information matrix diagonal it is only of use in practice when determined quickly. This is possible with a scalar weight matrix and where parallelogram designs and their extensions have been used.

The angular transformation

Fisher (1922) used the angular transformation in the study of binomial data. It is defined,

$$\pi_i = \sin^2 \rho_i$$

With this transformation the weight matrix becomes independent of the expected response. The matrix may therefore be computed once and for all at the beginning of the iterative procedure as it does not change from cycle to cycle. Standard errors of regression coefficients are therefore known in advance, and the design and allocation of experimental animals to groups may be adjusted to give any predetermined degree of accuracy. If the goodness-of-fit is unsatisfactory the assumptions on which these standard errors are based have been shown false and the observed variation must be used instead of the theoretical variation to judge the significance of regression coefficients. Their variances are enlarged by a heterogeneity factor, the goodness-of-fit χ^2 divided by its degrees of freedom (Finney, 1952a), and the t test of their significance is based on this number of degrees of freedom.

The weight matrix becomes,

$$\mathbf{W} = w\,\mathbf{diag}(n_i),$$

where $w = 1/820.7$ if ρ measured in degrees.

If the elements n_i are constant, i.e. if the group size is constant, say n, the weight matrix is given by,

$$\mathbf{W} = nw\mathbf{1}, \text{ a scalar matrix.}$$

If the weight matrix is scalar and the coordinate functions are orthogonal the information matrix is diagonal. The functions are orthogonal if,

$$\mathbf{X'X} = \mathbf{diag}(\cdots)$$

Granted a scalar weight matrix it may be cancelled from the estimation equations so the transformation matrix becomes,

$$\mathbf{T} = (\mathbf{X'X})^{-1}\mathbf{X'}$$

Claringbold, Biggers and Emmens (1953) showed that in factorial experiments where the total number of experimental animals exceeded about 300, the equation,

$\check{\mathbf{\xi}} = \mathbf{Tr}$, where $p_i = \sin^2 r_i$, and p_i is the l^{th} observed proportion,

gave estimates of a very similar value to the maximum likelihood estimates. Also it was suggested that in small samples the provisional estimates should be obtained with this equation.

Unequal group sizes.

With unequal group sizes the analysis in fully efficient form must be weighted. In this case the information matrix is non-diagonal and it must be inverted using more complicated procedures than forming reciprocals of the individual elements, as is done in the case of a diagonal matrix.

An alternative approximate procedure is available when the experimental groups suffer small losses during the course of an experiment and when the variance of this loss is known. Fisher (1925) has discussed an analogous problem in the derivation of a combined estimate from a number of estimates with faulty weights. It was concluded that if the variance of the false weights about their mean value was small, the combined estimate suffered little loss in efficiency if the estimates were weighted in terms of the average weight. In the present problem a similar solution is offered by the use of the average group size $\bar{n}$ to form the weight matrix. If the loss of experimental material is a random binomial variable with expectation about 1–2% it may be shown that the use of the average group size results in little additional loss of information. The loss is NQw where Q is the probability of loss of an experimental animal. In the type of experiments carried out in this laboratory a loss of 1% of experimental animals is unusually large.

Analysis of variance.

When the estimates of the regression coefficients have been obtained the standard analysis of variance may be carried out if the information matrix is diagonal. The total sum of squares of the working variate is

partitioned into a series of independent sums of squares corresponding with treatment effects.

Goodness-of-fit.

The weighted deviations from regression give a χ^2 test (Fisher, 1954) with $(N - g')$ degrees of freedom.

$$\text{i.e.} \qquad \chi^2 = \psi'\mathbf{W}\psi - \{\beta'\mathbf{I}\beta \text{ or } \psi'\mathbf{W}\varrho\}$$

Special case when $g' = N$.

Bailey (1951) has shown that when the number of regression coefficients requiring estimation is equal to the number of experimental points the observed values may replace the expected in the estimation equations. In this case the working variate degenerates into the observed or empirical angular response and a noniterative solution is carried out.

REFERENCES

Anscombe, F. J. (1954). Discussion of Fisher (1954). *Biometrics*, *10*, 141–144.

Bailey, N. T. J. (1951). Testing the solubility of maximum likelihood equations in the routine application of scoring methods. *Biometrics*, *7*, 268–274.

Bartlett, M. S. (1954). Discussion of Fisher (1954). *Biometrics*, *10*, 140–141.

Berkson, J. (1953). A statistically precise and relatively simple method of estimating the bioassay with quantal response, based on the logistic function. *J. Amer. Statist. Soc.*, *48*, 565–599.

Biggers, J. D. (1951). The calculation of the dose-response line in quantal assays with special reference to oestrogen assays by the Allen-Doisy technique. *J. Endocrinol.*, *8*, 169–178.

Biggers, J. D. & Claringbold, P. J. (1954). Optimum conditions for the local (intravaginal) action of oestrogens. *Aust. J. Biol.*, *7*, 118–139.

Biggers, J. D., Claringbold, P. J. & Emmens, C. W. (1954). A study of the variation in response of ovariectomized mice to the intravaginal and subcutaneous administration of oestrogens. *J. Endocrinol.*, *11*, 26–35.

Box, G. E. P. & Wilson, K. D. (1951). On the experimental attainment of optimum conditions. *J. R. Statist. Soc.*, *13B*, 1–45.

Campbell, R. C., Hancock, J. L. & Lord Rothschild (1953). Counting live and dead bull spermatoza. *J. Exper. Biol.*, *30*, 44–49.

Claringbold, P. J., Biggers, J. D. & Emmens, C. W. (1953). The angular transformation in quantal analysis. *Biometrics*, *9*, 467–484.

Claringbold, P. J. (1953). The effect of metabolic inhibitors on the local action of oestrone and oestradiol-3:17β on the vagina of ovariectomized mice. *Aust. J. Biol. Sci.*, *6*, 657–664.

Cochran, W. G. (1938). Some difficulties in the statistical analysis of replicated experiments. *Emp. J. Exp. Agric.*, *6*, 157–178.

Finney, D. J. (1952a). *Probit Analysis*. 2nd ed. Cambridge University Press.

Finney, D. J. (1952b). *Statistical method in bioassay*. Griffin, London.

Fisher, R. A. (1922). On the dominance ratio. *Proc. Roy. Soc. Edin.*, *42*, 321–341.

Fisher, R. A. (1925). Theory of statistical estimation. *Proc. Camb. phil. Soc., 22*, 700–725.

Fisher, R. A. (1934). Two new properties of mathematical likelihood. *Proc. Roy. Soc., 144A*, 285–307.

Fisher, R. A. (1946). A system of scoring linkage data with special reference to pied factors in mice. *Amer. Nat., 80*, 568–578.

Fisher, R. A. (1954). The analysis of variance with various binomial transformations. *Biometrics, 10*, 130–139.

Fisher, R. A. & Yates, F. (1953). *Statistical Tables.* 4th ed., Oliver and Boyd, Edinburgh.

Potter, C. & Gillham, E. M. (1946). Effects of atmospheric environment, before and after treatment, on the toxicity to insects of contact poisons. *Ann. Appl. Biol., 33*, 142–159.

Rao, C. R. (1952). *Advanced statistical methods in biometric research.* Wiley, New York

LETTER TO THE EDITOR

To the Editor of *Biometrics*

Sir,

Dr. Hamaker's excellent paper (Vol. 11, p. 257) contains one implication that I hope will be considered false by most biometricians: I can write thus bluntly, because it is clear that Dr. Hamaker would be pleased to find himself mistaken on this point. He appears to suggest that in biology, and especially in agriculture—in contrast to industry—an analysis of variance is considered an adequate summary of the statistical examination of an experiment without any tabulation of means and standard errors.

An experienced statistician may often be able to judge from inspection of an analysis of variance what the main features of the interpretation of the experiment are, especially of course if he has himself computed that analysis. This will not blind him to the fact that well-arranged tables of means and their standard errors are the most important summaries of numerical information from any experiment. The analysis of variance, far from being a more sophisticated presentation of the information, is usually no more than a scaffolding needed in preliminary study of the data: it does not require inclusion in the ultimate report, since its duty is accomplished when it has indicated what tables of means need to be discussed and has provided estimates of standard errors. These functions are made clear in Hamaker's §8.

A contrary impression may sometimes be given by text-books and papers expository of statistical techniques. Naturally these give special attention to matters that may be unfamiliar to their readers, amongst which often are details of calculations of the analysis of variance for new experimental situations. Economy of space may prevent the authors from discussing how to prepare and present, for each set of data, good summary tables or diagrams, procedures with which they may assume—not always correctly—their readers to be familiar. Perhaps authors of text-books could usefully point out that while attempting to teach statistical techniques they cannot always be illustrating how best to present interpretative reports.

There are of course exceptions to my suggestion that the analysis of variance of an experiment does not need to be reported. Apart from its expository value, for example, it may give information on components of variance important to any discussion of the efficiency of the design and the possibility of improving it for subsequent work.

It would be unfortunate, however, if novices in biometric practice were to imagine that abandonment of tables of means, the natural and common-sense summaries of experimental data, in favour of tables of analyses of variance was an advance in statistical sophistication at which they should aim.

I do not question the appropriateness of Dr. Hamaker's tabulations for their purpose; I entirely agree that in other fields of application different styles of presentation may be needed, and that the style needs to be adapted both to the character of the data and to the statistical experience of the reader; I should deplore any tendency of statisticians to expect readers of their reports on experiments to comprehend 'sums of squares, degrees of freedom, and mean squares' *instead of* means of observations.

Yours truly,

D. J. FINNEY
Department of Statistics
Marischal College,
University of Aberdeen,
Aberdeen, Scotland

25 October, 1955

QUERIES

GEORGE W. SNEDECOR, *Editor*

118 **QUERY**: The following problem occurs frequently in some of my studies and I would very much appreciate your opinion as to whether the following method of solution is correct.

Samples of unequal numbers of fish have been collected from 5 different locations in a lake. A one-way classification of analysis of variance indicates that their length is significantly different at the .01 level. Apparently it is the average size that is different since Bartlett's test on the variances is not significant. The problem is to establish whether any particular location or locations are reponsible for the difference while the others can be considered to belong to the same population. I believe we could find this out by calculating the fiducial limits of the means. The variance of the population would be the mean square within locations but I am uncertain what to divide this by for the calculation of the 95% semi-interval.

Should the mean square be divided by k_0 as per the formula on Page 234 of Snedecor's 4th edition?

This method would provide fiducial limits for all means in one calculation. Sometimes the numbers of fish in each sample vary greatly e.g., 35, 23, 7, 16, 41 = 122 fish for 5 locations. In this case would it be more correct to calculate a fiducial limit for each mean separately using the common mean square within population and their respective n's?

ANSWER: If a 95% fiducial interval is wanted, *separately* for the ith "location mean" μ_i this can, as you say, be computed from the usual formula as

$$\bar{x}_i \pm t_\nu(5\%) \cdot s\sqrt{k_i^{-}} \qquad (1)$$

where s^2 is the 'within location' mean square, ν its degrees of freedom and $t_\nu(5\%)$ the corresponding double tail % point of t, while k_i is the sample size for the ith location mean $\bar{x}_i$.

If a fiducial interval is wanted for each of the n location means μ_i then the above computation can of course be carried out for all locations. Computational labor may be saved by using an 'average' sample size k_0 in place of the separate k_i , but this appears hardly worth while. The question of *what* average of the k_i may be used arises. The answer depends on what properties it is desired the average fiducial interval should have. The formula for the average group size k_0 given by

Snedecor on p. 231 arises in the estimation of components of variance and it is difficult to conceive of a realistic reason why it should be used here. If it is desired that the average confidence coefficient should still be 0.95 this could be (approximately) achieved by using for the average sample size the value

$$\sqrt{k} = \frac{1}{n} \sum_{i=1}^{n} \sqrt{k_i} \tag{2}$$

An analogous problem arises when you are concerned with setting up a confidence interval for the *difference* between two particular location means $\mu_i - \mu_j$ which is given by the familiar formula

$$(\bar{x}_i - \bar{x}_j) + t_\nu(5\%) \cdot s\sqrt{\frac{1}{k_i} + \frac{1}{k_j}} \tag{3}$$

Again, the problem of averaging may be posed here.

Finally the question arises as to how the fiducial intervals be used to decide the question as to which of the locations differ with regard to the fish-length. Presumably your idea was to call any two locations (i, j) with mean fish lengths μ_i , μ_j for which the confidence interval (3) does not include the value 0 different from one another. With such a procedure you would be making $n(n-1)/2$ decisions about the $n(n-1)/2$ differences and the question of your 'error-rate', i.e. your frequency of wrong statements, now becomes a little more complex than in the case of a single fiducial interval for a single pair of means.

A solution to your problem is given by Henry Scheffé in "A method for judging all contrasts in the analysis of variance," Biometrika 40:46–49 and 57–62 (1953). You will get a 5% overall experiment-wise error rate if you replace in (3) $t_\nu(5\%)$ by $\sqrt{(n-1)F}$. F is the 5% point of F for $n-1$ numerator degrees of freedom and ν denominator degrees of freedom.

H. O. HARTLEY

CHANGE IN THE EDITORSHIP OF BIOMETRICS

With the December, 1955 issue of *Biometrics*, Miss Gertrude Cox will terminate her period of service as editor of *Biometrics*, a post which she has held since the first number came out, under the title *Biometrics Bulletin*, in February, 1945. On several occasions during recent years, Miss Cox has asked to be relieved of the editorship owing to the increasing pressure of her many other commitments in biometric activities. These requests were handled either by persuading Miss Cox to continue in the post, or in some cases, I fear, by pretending that they had not been heard. When the request was renewed in 1954, however, it was felt that the Society had imposed too long on Miss Cox's public-spiritedness, and that steps should be taken to seek a successor. A committee consisting of F. Yates (Chairman), D. Mainland and A. Linder was appointed to make a recommendation about a successor. After careful consideration of a number of possible nominees and sites, the committee unanimously recommended Dr. John W. Hopkins of the Division of Applied Biology, National Research Council, Ottawa, Canada, who kindly consented to take on the post after due approval by the Council. During the present year Dr. Hopkins has been serving as Associate Editor in order that a smooth and orderly transition of the work can be made. With the March 1956 issue he will assume full responsibility as editor.

I do not know whether members realize what a great debt they owe to Miss Cox. From an initial issue of 12 pages, she has built up *Biometrics* into a leading journal in its field. Her refereeing policy has been helpful and considerate to authors, while maintaining standards of high quality. She has devoted a great deal of thought and effort to obtaining expository papers and simple examples of the newer techniques that would be easily understood by biologists. Some of these efforts were unrewarded, for the number of competent people who are willing to prepare expository papers, or to deliver them after having agreed, is distressingly small. Many of the special issues of *Biometrics* that have been in wide demand were planned and stimulated by Miss Cox. She has kept the publication on a sound financial basis through the troublesome early years, despite some disappointments in securing financial aid that had been anticipated. The steadily growing requests for complete sets of the journal from libraries, institutions and individuals, is a testimony to the high regard in which it is held. I want to

take this opportunity on behalf of the members of extending to her our warmest thanks.

Thanks are also due to Mrs. Sarah Porter Carroll of the Institute of Statistics, North Carolina State College, who has served diligently and competently as managing editor of *Biometrics*, and has relieved Miss Cox of much of the time-consuming labor that is unavoidable in editing a journal. North Carolina State College has contributed generously by making available secretarial help, space and facilities.

Dr. Hopkins, who is Chairman of the Finance Committee and has served on the Council, will be well-known to most members. I am sure that he can count on the full cooperation of members in maintaining the high quality that has been set.

W. G. COCHRAN
President

ABSTRACTS

Biometric Society Meetings (WNAR), Pasadena, June 23-24, 1955

325 PAUL E. FIELDS (School of Fisheries, University of Washington). **Factorial Designs and the Guidance of Downtsream Migrant Salmon and Steelhead Trout.**

This study is a part of the Columbia River Fisheries Engineering Investigation and Research Program sponsored by the North Pacific Division, Corps of Engineers. Probably because of difficulties encountered in rearing anadromous salmon, there was but little basic information about their sensory abilities and their behavior patterns when it suddenly became necessary to find some effective guiding stimulus. The avoiding response to light seemed to offer the most promise of success.

In the first series of factorial experiments, the reactions of a total of 90 different groups of 25 one year old silver salmon, each given four trials, was obtained to a light barrier with three angles and three light intensities, in water of four velocities and three depths. In general, the number of fish entering the lighted area was significantly reduced as the angle of the barrier was made smaller, the intensity of the light was increased, and the velocity was decreased. The F for depth was not significant. In a second experiment, the responses of a total of 72 groups of steelhead trout, chinook and silver salmon were compared on two barrier angles, two light intensities, and two water velocities. The findings of the previous experiment were confirmed and the range extended. In addition, a species difference was established with steelhead trout being the most sensitive to light. In a third experiment, the reactions of a total of 48 groups of 50 each of steelhead trout, chinook and silver salmon were compared with respect to chain barriers with two angles, two different spacings and two water velocities. The only significant F was between species.

As the success in guiding has increased, the normality of the data has decreased, making the application of the usual parametric methods more questionable.

326 D. G. CHAPMAN AND R. PYKE. (University of Washington). **The Statistical Theory of Some Migration Population Models.***

The problem considered is that of estimating the parameters associated with the migration of individuals between two areas (A_1 and A_2). The populations studied are comprised of two classes, the X-type and

*Work done partly under the sponsorship of the U. S. Office of Naval Research.

the Y-type individuals. Let $X_i(Y_i)$ be the number of X-type (Y-type) individuals in A_i before migration. We assume that these parameters are known and that $P_1 \neq P_2$ where $P_i = X_i/N_i$ and $N_i = X_i + Y_i$, $(i = 1, 2)$. Define M_{xi} as the number of X-type individuals migrating from A_i to A_{3-i}, and define M_{yi} similarly. Set $M_i = M_{xi} + M_{yi}$. The following models have been studied.

Model I: Assume

(a) migration occurs in one direction only, from A_1 to A_2, say;
(b) M_{x1} is a random variable, distributed as $b(M_{x1} : M_1, P_1)$
(c) a sample of size n_2 is taken in A_2 after migration in which X_2 X-type individuals are observed.

The estimator of M_1,

$$\hat{M}_1(x_2) = \begin{cases} N_1 & \text{for } x_2 \text{ between } j \text{ and } n_2(X_1 + X_2)(N_1 + N_2)^{-1} \\ N_2 \dfrac{n_2P_2 - x_2}{x_2 - n_2P_1} & \text{for } x_2 \text{ between } n_2P_2 \text{ and } n_2(X_1 + X_2)(N_1 + N_2)^{-1} \\ 0 & \text{otherwise} \end{cases}$$

where j is 0 or according as $P_1 - P_2$ is negative or positive, is derived, studied and the conditions for its reliability outlined. Approximate formulae for the expectation and variance of it are given. It is shown that for large parameter values, x_2 is approximately distributed as

$$b\left(x_2 : n_2, \frac{X_2 + MP_1}{N_2 + M}\right)$$

Model II: Assume (a) and

(d) simultaneous samples of size n_i are taken in A_i after migration in which x_i X-type individuals are observed $(i = 1, 2)$.

The M.L. estimators of M_1 and M_{x1}, for $z_i = x_i/n_i$, are

$$\tilde{M}_1(z_1, z_2) = \frac{N_2(P_2 - z_2) + N_1(P_1 - z_1)}{z_2 - z_1}$$

$$\tilde{M}_{x1}(z_1, z_2) = \frac{z_1N_2(P_2 - z_2) + N_1(P_1 - z_1)}{z_2 - z_1}$$

of which the asymptotic properties are studied.

Model III: Assume (a) and

(e) the migrants are distinguishable from the residents,

(f) a sample of size n is taken in A_2 after migration in which r_i X-type and s_i Y-type individuals are observed, the subscript denoting the area in which the individuals were before migration

The M.L. estimators of M and M are

$$\overline{M}_1(r_i\ ,\ s_i) = \frac{(r_1 + s_1)N_2}{r_2 + s_2}\ , \qquad \overline{M}_{21}(r_i\ ,\ s_i) = \frac{r_1 N_2}{r_2 + s_2}\ ,$$

of which the asymptotic properties are studied.

Model IV: Assume (d) and

(g) migration occurs in both directions.

In this case, the same estimators are obtained as in Model II, except that now negative values of these functions make sense.

327 JOSHUA L. BAILY, Jr., Sc.D. (San Diego, California). **Variation of the *Pecten Gibbus* Complex.**

This is a repetition of Davenport's "Quantitative Studies on the Evolution of *Pecten*" made about half a century ago, to see what changes, if any, have taken place in the meantime.

Pecten is a bivalve mollusc, and like other bivalve molluscs has two valves which are organically right and left. But the species in this investigation when at rest, and also when swimming have changed their ancestral orientation, the primitively right and left valves having become in both cases the lower and upper valves respectively, and the functional right and left halves being organically anterior and posterior. The question then arises as to whether in general functional considerations are more influential than organic relationships in determining correlations.

Davenport also concluded that species from the Pacific coast are in general more variable than closely related species from the Atlantic coast. This differential variability he concluded could be attributed to the differences in the geological history of the two coasts. But Davenport's conclusions were based upon measurements of dimensions, and the difference in variability in size might be more simply explained by assuming a greater number of age groups represented in the Pacific series of specimens. Coefficients of variation of ratios might conceivably have a different result, since they would indicate variability of shape rather than of size.

R. F. TATE AND R. L. GOEN. (University of Washington).

328 Minimum Variance Unbiased Estimation for a Truncated Poisson Parameter.

A problem of current interest, especially in public health work, is that of estimating the parameter of a Poisson distribution which is singly truncated on the left. An MVUE is found for the general case, with simple expressions resulting for the two important special cases of truncation away from zero, zero and one. Results proceed from considerations involving sufficient statistics.

HERBERT D. KIMMEL. (University of Southern California).

329 The Reliability of Categorical Qualitative Judgments.

While the problem of estimating the reliability of quantitative data such as test scores or qualitative data which may be artificially quantified along one qualitative dimension has been met adequately by application of one of several methods based on score-variations, no generally applicable method has been devised to estimate the reliability of a qualitative rating or sorting schema which cannot be artificially quantified.

This paper proposes a new method for estimating reliability in such situations which is based on the proportion of agreement obtained among several judges and which takes into account the proportion of agreement which would be expected to obtain by chance alone. The reliability estimate obtained is logically analogous to internal consistency type measures, on the assumption that each individual judge acts as a separate item in a test. The method gives values ranging between zero and unity.

In addition to providing a reliability estimate for the whole schema, the method may be used to obtain separate reliability estimates for the separate qualitative categories. It should be noted, however, that these separate category-estimates are somewhat dependent upon the number of times the category was used. Also, the method is not recommended in situations which may be artificially quantified with reasonable justification.

DAVID A. GRANT. (University of Wisconsin).

330 Statistical Tests in the Comparison of Curves (by Means of Orthogonal Components of Trend).

This paper extends the well-known Alexander Trend Analysis procedure in two respects. The Alexander procedure applies where

there are a series of scores, obtained by repeated trials on the same *Ss*, in two or more groups. It provides for comparison of the groups in terms of: (a) their mean differences; (b) differences in the linear components of the group trends; and (c) differences in the pooled higher-order components of the group trends. The present procedure is more analytic in that: first, the groups may be compared separately in terms quadratic, cubic, and any further orthogonal components of the trends; and secondly, if the groups form an orthogonal array, *e.g.*, rows and columns, the row, column and interaction variation may be examined separately for linear, quadratic, cubic, etc. differences.

The tests are obtained by constructing covariance terms by means of the orthogonal polynomials. Using Cochran's theorem, it is easily shown that, with a mathematical model, linear in the orthogonal polynomial components, with normal, random, and equal error variation, the separate component tests conform to the F distribution. A routine method of calculation of the sums of squares has been worked out with suitable checking procedures.

The procedure is limited to cases where the intervals between trials or levels of the corresponding independent variable are equal on a linear, logarithmic or similar scale. It has proved most valuable in our laboratory for comparing experimental curves separately with respect to slope, curvature, inflections, and the like. It is not particularly efficient when the curves are expected to follow exponential or other transcendental functions.

331 PHILIP R. MERRIFIELD. (University of Southern California).
Quantification of Ordering Behavior.

Ordering behavior occurs in several contexts. It is defined here as the process of arranging objects or situations, or verbal definitions thereof, in the order most appropriate with reference to a criterion. The criterion may or may not be stated explicitly, but in general it has the characteristics of (a) a time continuum, (b) a spatial arrangement, or (c) an hierarchical system. These three are contextual cases of what might be labelled as "logical" arrays.

Two major alternative hypotheses as to the nature of the ordering process are entertained. Under the first, it is suggested that the examinee treats the set of stimuli as a whole, transforming the entirety into a new array by a process that is primarily "unitary;" subhypotheses deal with minimizing the "error space" in what correspond to the two-dimensional and three-dimensional cases. Under both subhypotheses, double and triple interaction effects are considered.

Under the second major hypothesis, it is suggested that the examinee deals separately with the individual elements in the problem, or at most considers the individual elements in pairs.

On the premise that scoring systems based on these hypotheses will disclose differences in total scores sufficient to support a decision in favor of one or the other, seven scoring systems are derived.

Selected results from a factor-analytic investigation of planning abilities, carried out by R. M. Berger, J. P. Guilford, and P. R. Christiansen, and from a smaller separate study carried out by the writer, are discussed with reference to the hypotheses concerning the nature of the ordering process.

332 J. A. GENGERELLI. (University of California, Los Angeles). **A Method of Constellation Analysis.**

The intent of the method is to determine whether an assemblage of objects is comprised of mutually exclusive sub-classes or whether it constitutes a single continuum. The problem arises in a variety of contexts, viz.: the question of (a) physical types, (b) personality types, (c) psychiatric nosologies, (d) psychological factors, (e) cultures.

The central concepts are those of "neighbors," "neighborhood," "distance," and "constellation." A constellation is defined and said to exist if a set of objects are mutually neighborly; two or more objects are mutually neighborly when all are at no greater distance one from another than a certain critical value s. The critical value s for any given universe of discourse is defined as the average of the distances separating every pair of objects constituting that universe. In certain contexts, however, a constellation is not determined by distance but by the fact of concomitance. Thus, a set of objects is said to be a constellation if the objects are concomitant, i.e., all occur if any one occurs.

A method for isolating constellations is described and applied to representative problems.

333 HARRY H. HARMAN. (The Rand Corporation). **Some Observations on Factor Analysis.**

This paper is non-mathematical and expository in character. A brief review is given of the origin and growth of factor analysis, delineating (1) its use in the formulation of psychological theories of human abilities and traits, and (2) its consideration as a branch of general statistical technique. The question is raised regarding the basis of choice of a particular factor solution out of the infinite possi-

bilities that arise in reducing a given matrix of correlations. Preferred types of solutions are enumerated, with an indication of the extent to which the two general principles—statistical simplicity and psychological significance—affect each type. Some of the problems that have plagued factor analysts during the past half century are pointed up. Finally, an appraisal of the present status of the subject and a prognosis for the future is made.

334 J. W. FRICK. (University of Southern California). **The Effect of Varied Interpolated Stimuli upon the Time Order Error.**

It was hypothesized that varied stimuli, interpolated between a constant stimulus and certain variable stimuli, would have a differential effect upon the judgment of the variable stimuli as compared with the constant. It was further hypothesized that, as the interpolated stimuli approached the constant in size, reinforcement of the latter would occur, resulting in a lesser number of negative TOE's.

Twenty *S*'s made a total of 1600 judgments of 4 randomly-presented variable stimuli in comparison with a constant stimulus, under four conditions of interpolated stimuli. All stimuli were black lines projected tachistoscopically on a constantly-lighted white screen, and varied in projected length from 18.8 to 21.2 inches. The constant was 20 inches in length.

As expected, the number of negative TOE's exceeded the number of positive errors, but were reduced to a frequency less than that to be expected by chance. An analysis of variance disclosed no significant differences between judgments made under the varying stimulus conditions. This may indicate that all interpolations had a reinforcing effect upon the constant stimulus, since the largest interpolated stimuli varied only plus-or-minus 2 j.n.d.'s (1.2 inches) from the constant. (The author is indebted to Dr. Harvey F. Dingman for the analysis of the data.)

335 DR. LLOYD A. LIDER. (University of California at Davis). **A Group of Long-term, Perennial and Non-replicated Root-stock Trials.**

Data on production, fruit quality and vigor have been taken on a group of cooperative grape rootstock trials established in the grape growing areas of the coastal counties of California. These trials, using experimental phylloxera resistant rootstocks, were set up over the

last twenty years, and annual measurements have since been taken. From seven to ten rootstocks appear in each non-replicated trial. They have been planted under a wide range of environmental conditions and have represented most of the commercial scion varieties of the area. From the data on hand it is possible to draw certain conclusions concerning the influence of the rootstocks on the behavior of the scion varieties. An interpretation of these data are presented as well as general conclusions concerning the usefulness of and difficulties presented by a study of this nature.

CLYDE STORMONT. (University of California at Davis).

336 Estimates of Frequencies of B Alleles in Three Breeds of Dairy Cattle.

Computational and theoretical difficulties in estimating the frequencies of even such well-adapted genes as those controlling blood groups trace to the frequent operation of extensive allelic series and the necessity of taking family relationship into account. Our model here is the 100 or more alleles which control the exceedingly complex *B* system of bovine blood groups—a system which has maximum utility in present and projected application along such diverse lines as medicolegal problems, evolution theory, clinical genetics, animal breeding plans, immunology, genetic theory and so on. It has been known for a number of years that many of these *B* alleles are in principle almost breed specific but precise statements as to their frequencies within breeds have not been made. Marked differences in frequencies between lines within breeds pose a problem of obtaining samples that are representative of the various breeds. The computational problem could be reduced to its simplest form by developing serological reagents capable of differentiating all of the 5,050 or more phenotypes in the *B* system. Considerable progress towards this goal has been made and consequently, efficient estimates with small errors due to misclassification of some of the phenotypes can be obtained by simply computing the ratios of the actual allele counts. Assuming there are no sex differences in regard to frequencies of *B* alleles, the sampling of bulls in various semen producing businesses throughout the United States would seem to provide one of the best methods of obtaining samples that are representative of the various dairy breeds. "Preliminary" estimates of the frequencies of *B* alleles in Guernsey, Jersey and Holstein-Friesian breeds in U. S. are based on samples of 200 Holstein-Friesian, 80 Jersey and 80 Guernsey bulls.

337 HERMAN RUBIN. (Stanford University and University of California at Berkeley). **Axiomatization of Genetics.**

Given a theory of genetics (such as Mendelian inheritance without mutations), it is desired to construct a mathematical system which has the properties the geneticist ascribes to his theory. One wishes to list a number of mathematical entities and a number of axioms connecing them: also to call the mathematical model an axiomatization of genetics, it is necessary to correlate genetic terms with those mathematical entities, called *primitive notions*, which are not explicitly defined. In addition, genetic terms might correspond to defined entities; for example, if the relation *immediate ancestor* is a primitive notion, the relation *ancestor* could be a defined notion. In most genetic models, especially those on a cellular basis, a large number of primitive notions is required. For example, in Mendelian inheritance, *haploid cell, diploid cell, genetic map, genetic description*, and the relations of *mitosia, meiosis*, and *fertilization* would normally be some of the primitive notions. Typical examples of axioms are *A genetic map is a subset of Euclidean two dimensional space with the x-coordinate an integer*, and *If A is related to B by mitosis, the genetic descriptions of A and B are identical.* An example of a statement which would not belong to a genetic axiom is *During mitosis, the new chromosomes are drawn to their respective nuclei.*

Australasian Symposium on Biometry, University of Melbourne Australia, Monday, 22nd August, 2:15 p.m.

338 R. T. LESLIE. (University of Melbourne). **A Statistical Approach to the Physiological Problem of Thresholds.**

For a stimulus of given intensity I, a certain change ΔI is necessary for the difference to be detectable; the classical Weber-Fechner law is to the effect that within a certain range $\Delta I/1 =$ constant.

It is known that a stimulus is translated into electrical pulses in the receptor nerve, the frequency of the pulses being related to the stimulus intensity. To distinguish two stimuli as of different intensity the discriminating organ must compare two pulse rates, known to be each subject to random fluctuations. Assuming storage of information, the problem is analagous to the comparison of a pair of samples, where sampling is over (n) units of time, and discrimination should improve

in proportion to $1/\sqrt{n}$. A similar argument may be applied to sampling of stimulated area (visual and tactile stimuli), discrimination again improving with increase in area.

To explain why $\Delta I \neq 0$ for sufficiently protracted or intensive stimuli ($n \to \infty$) random noise in the central nervous system may be assumed, the noise adding a constant and irreducible amount to the variance of the difference between the impulse rates.

Some further predictions from this theory have been verified experimentally.

339 E. J. WILLIAMS. (Division of Mathematical Statistics, C.S.I.R.O., Melbourne). **Sidelights of Sampling Surveys.**

Some of the problems arising in the conduct of sampling surveys and the interpretation of their results are discussed. A survey seldom works out in the field as originally planned.

Some experiences with the New Guinea Census of Native Agriculture are described. How should one treat a sample village which is found to have—

(i) disappeared,
(ii) migrated to another district, or
(iii) split up into several villages?

The effects of methods of sampling and of ascertainment of attributes on the method of estimation is considered. Cases arise where the probability of inclusion of an item is—

(i) proportional to size,
(ii) inversely proportional to size,
(iii) modified by the method of sampling adopted.

Each case requires a different form of estimate.

The problem of bias in estimates from small samples is important for sampling surveys. Methods of reducing bias are discussed.

340 G. S. WATSON. (Australian National University, Canberra). **Missing and Mixed-Up Values in Contingency Tables.**

In the analysis of variance, the problems of missing and mixed-up plots have well-known solutions. However the same problems may arise in the analysis of contingency tables. It is shown in this paper that they may be solved by an application of the method of maximum likelihood.

341 P. J. CLARINGBOLD. (University of Sydney). **Discriminant Analysis in the Interpretation of Semi-quantal Data.**

Intermediate between quantal responses and the fully quantitative or graded responses are responses which may be described by the term semi-quantal. These may simply be regarded as a generalization of the quantal response to the case where more than two response classes are recognised. In the past, data of this type from insecticidal studies were always reduced to a quantal form by grouping of response classes. This was carried out because of mathematical difficulties associated with the extension of probit analysis to semi-quantal response. In his "Statistical Methods" Fisher shows how discriminant analysis may be employed to derive scores for various types of response. The method described is extended in this paper and illustrated by an example taken from studies on the action of oestrogens.

342 (Mrs.) G. L. RICHARDSON and F. E. BINET. (University of Melbourne). **Discriminant Analysis on Species of the Genus *Murravia* (Brachiopoda, Tertiary and Recent).**

Specimens of both *M. triangularis* and *M. lenticularis*, obtained from different localities (all restricted to the "lower Aldingan" of Tate), exhibit little or no variation in specific characters. Statistical analysis on samples of *M. triangularis* shows that there is no evidence against over-all homogeneity in these samples.

M. catinuliformis shows a considerable range of variation both within and between different collections which are derived from deposits over a wide stratigraphic range (Janjukian to Cheltenhamian Stages). These collections are empirically divided into three groups which, it is suggested, may form a basis for future specific or sub-specific distinction.

A statistical analysis is based on the logarithms of (1) breadth, (2) distance from the posterior point of the pedicle valve to the line of greatest breadth, (3) distance from the anterior tip of the pedicle valve to the line of greatest breadth.

A size factor is first defined as the linear combination with coefficients equal to the reciprocals of the estimated standard deviations. General linear combinations are then formed, and from those which appear to be uncorrelated with the size factor, two are chosen, one maximizing and the other minimizing the variation between collections relative to that between individuals. The variation is significant in both directions. These optimal combinations are significantly better discriminators than those derived from indices suggested by the morphology of the Genus.

NOTE

During the past few years I have been repeating the investigation of variability of *Pecten* made by Davenport half a century ago.

One of the valuable features of this work was the extensive bibliographies, with references to parallel work in other fields. The interest and value of the report of my work would be greatly enhanced if I could add to it references to similar investigations carried on since Davenport's work was published.

Davenport found the Pectens of the Pacific coast to be more variable than those of the Atlantic. I agree with this belief, but I would like to know if it holds true for the other forms of life, either animal or vegetable. Are Pacific coast species always more variable than geminate species from the Atlantic slope?

He published a list of investigations of the correlation between the right and left paired structures of bilaterally symmetrical organisms. I would like references to work since 1905 in which such correlations are reported.

Joshua L. Baily, Jr.
4435 Ampudia Street
San Diego 3, California

ANNOUNCEMENT

Research proposals directed to the Division of Biological and Medical Sciences of the National Science Foundation of the U.S.A. will be received at any time. The proposals on research projects to begin in June or September 1956, will be reviewed during March. These proposals should be received by the Foundation prior to February 1, 1956.

Projects in the areas of anthropology, human ecology, functional archaeology, experimental social psychology, and demography are included in the Division's program.

NEWS OF MEMBERS

Dr. R. Lowell Wine, who received his B.A. degree at Bridgewater College, major in Mathematics, and his M.A. degree at the University of Virginia, major in Mathematics, and his Ph.D. in Statistics at the Virginia Polytechnic Institute, has joined the staff of the Department of Statistics of the Virginia Polytechnic Institute as Associate Professor. He has previously taught at the University of Virginia, Amherst College, University of Oklahoma, and Washington and Lee University. He was

a student in Statistics at the University of Michigan for two years prior to coming to Virginia Polytechnic Institute. Dr. Wine will do both teaching and research.

Dr. Rudolf J. Freund has joined the staff of the Department of Statistics as Associate Professor at the Virginia Polytechnic Institute. Dr. Freund did his undergraduate work at North Carolina State College, received his Master's degree in Statistics and Economics at the University of Chicago, and his Ph.D. degree in Statistics at North Carolina State College. Dr. Freund will do both teaching and research.

Helen Bozivich, research associate in the Statistical Laboratory of Iowa State College, has accepted a position at Purdue University as assistant professor in its Statistical Laboratory beginning July 1, 1955. She was awarded the degree of Doctor of Philosophy in statistics by Iowa State College in June, based partly on a dissertation, "Power of test procedures for certain incompletely specified random and mixed models."

John F. Hofmann, chief statistician of the Naval Ordnance Laboratory at Corona, California, was granted a Ph.D. degree in statistics by Iowa State College June 1955. His dissertation concerned "Life testing in controlled environmental conditions."

Bernard Ostle has been appointed Professor of Mathematics and Statistics at Montana State College.

In December 1954, the University of Melbourne created a full chair in Statistics and appointed Associate Professor **M. H. Belz** as the first Professor. This is the first full chair in Statistics at a teaching University in Australia. **Dr. G. S. Watson** of the University of Melbourne has accepted an appointment as Senior Research Fellow in the Department of Statistics at the Australian University in Canberra, effective in March of this year. The new Senior Lecturer at Melbourne University will be **Dr. H. A. David** of the Department of Mathematical Statistics, C.S.I.R.O.

John W. A. Brant, formerly Agricultural Officer of the Food and Agriculture Organization of the United Nations (1953–1955), now Specialist of the Universidad de Guayaquil y Universidad de Idaho en Programa Cooperativo para el Progreso de las Ciencias Agropecuarias, has been honored November Eighteenth by nomination to Professor, Facultad de Agronomia y Veterinaria during a Sesion Solemne de la Facultad de Agronomia y Veterinaria, Universidad e Guayaquil, Guayaquil, Ecuador—Octogésimo octavo aniversario de su Fundación. He has launched a research program in poultry nutrition, which is to be continued concurrently with research programs in animal physiology and genetics.

INTERNATIONAL BIOMETRIC SYMPOSIUM ON "THE ROLE OF BIOMETRIC TECHNIQUES IN BIOLOGICAL RESEARCH"

General Proceedings

Instituto de Educacao Carlos Gomes, Campinas, Brazil

July 4–9, 1955

The second International Biometric Symposium, on the role of Biometric Techniques in Biological Research, met under the sponsorship of the Biometric Society, a section of the International Union of Biological Sciences, and under the auspices of the University of São Paulo represented by its Seminario de Estatistica. The Symposium was convened by the President of the Society, Professor W. G. Cochran, shortly after 10 a.m. on July 4. He introduced the Secretary of Agriculture for the State of São Paulo, Dr. R. Cruz Martins, who welcomed the Symposium to Brazil and Campinas and wished it success in its deliberations. In reply, President Cochran thanked Dr. Cruz Martins for his country's hospitality and good wishes. He then delivered his presidential address on the 1954 poliomyelitis trial in the United States, as an illustration of the critical role played by biometry in solving a major public health problem.

Following his address, President Cochran called upon Secretary Bliss for a summary of recent Society activities. The Secretary noted that proceedings of the Third International Biometric Conference of the Society, at Bellagio, Italy, in September, 1953, had appeared the following December in Biometrics, which has also published a number of the papers presented at the Conference. Under the editorship of Professor G. M. Cox, Biometrics had attained world-wide repute, but after 11 years, she had asked to be relieved of this post. The Society was fortunate in having obtained Dr. John W. Hopkins of Ottawa, Canada, as her successor, beginning with Volume 12.

In sponsoring the present Symposium, the Society was also acting in its capacity of Biometric Section of the International Union of Biological Sciences. The Union had allotted $2000 toward the travelling expenses of principal participants and $500 towards the publication of our proceedings. All papers read during the Symposium may appear in abstract in Biometrics and a number of them in full. At the XII General Assembly of the IUBS in Rome this past April, the section was represented by Drs. L. L. Cavalli-Sforza and A. Vessereau. Again with financial assistance from the IUBS, the Society (or Section) was

sponsoring an international Biometric Seminar to be held in Varenna on Lake Como, Italy, on September 7 to 23 of this year.

He noted further that many biometric meetings have been sponsored by national and Regional subdivisions of the Society in all parts of the world, so that in 1954 alone more than 20 of these had been reported. The Society now has a membership of 1300 with eight organized Regions and Regions authorized in Japan and also in Brazil when a sufficient number of members has been enrolled. The Brazilian members hoped to reach this goal before the end of the present Symposium. An organization meeting had been scheduled on the program. Following announcements by Dr. C. G. Fraga, the morning session was adjourned.

The first scientific session on Monday afternoon concerned biometrical genetics with the papers listed in the Scientific Program. A group of geneticists at the afternoon program met again in the evening to hear an address by H. Kalmus.

Experimental design was the subject of both sessions on Tuesday, July 5, the papers in the morning dealing especially with perennial crops. At 5 p.m. the participants in the Symposium were received in a pleasant reception by the Mayor of Campinas. In the evening, officials of the Biometric Society met with Brazilian scientists to discuss the formation of a Brazilian Region. At the end of a lively discussion, a committee was named to prepare a tentative set of Regional by-laws.

The following morning the Symposium moved by bus to Piracicaba for the day, a distance of 60 kilometers, where it continued with a panel discussion on experimental designs for perennial crops at the Escola Superior de Agricultura "Luiz de Queiroz", the Agricultural College of the University of São Paulo. Following a delightful luncheon of typical Brazilian dishes as guests of Professor Brieger's Department of Genetics, members of the Symposium visited the School and its associated Experiment Station and then were taken on a tour to a nearby sugar mill, to a paper factory, which used the bagasse from the sugar cane as raw material, to the Sugar Experiment Station, where its research program was summarized briefly, and finally to a social club in Piracicaba for coffee, refreshments and music before returning to Campinas by bus.

The morning session on Thursday, July 7, concerned the statistics of animal feeding experiments. The committee on by-laws for the Brazilian Region met at lunch and continued their work through the afternoon. Sampling techniques was the subject of the afternoon session of the Symposium. Members of the Council and other officers of the Society attending the Symposium met at dinner for informal discussions. In the evening, the Tenis Clube of Campinas entertained

the members of the Symposium at a social gathering, highlighted by a superb exhibition of native dances by members of the Club.

On Friday, July 8, the morning was devoted to a visit to the Instituto Agronômico which has its headquarters in Campinas. Founded in 1881 and now supported entirely by funds from the State of São Paulo, it is the oldest agricultural research institution in Latin America. In addition to its main Experiment Station of 2000 acres, located near Campinas, the Institute has 19 branch Stations in various ecological areas of the State and 31 technical sections grouped into four divisions of agronomy, biology, soils and technology, and experiment stations. The afternoon session presented three papers and a panel discussion on bioassay. Following the panel discussion, the Brazilian members of the Society completed the formation of the Brazilian Region. In the evening the members of the Symposium were entertained at a barbecue and dance by the Sociedade Hipica at its ranch near Campinas.

The last scientific session of the Symposium, on the morning of July 9, concerned medical statistics. At the final meeting the following resolutions were adopted unanimously.

"The officers and members of the Biometric Society meeting at Campinas, 4–9 July 1955 extend their sincere appreciation to the Committee on Arrangements: Chairman, F. G. Brieger, and his committee: P. Mello Freire, F. Pimentel Gomes, A. Groszmann, A. M. Penha, W. L. Stevens and C. G. Fraga Junior, Executive Secretary. Our special regards are due to C. G. Fraga Junior for his patience and kind consideration of our individual and collective problems.

It is resolved that our thanks be expressed to Director C. A. Krug and his staff at the Instituto Agronômico; to Director E. R. Nobre, Escola Superior de Agricultura "Luiz de Queiroz"; and to Geneticist F. G. Brieger and his staff. The field trips provided an opportunity for scientific activities combined with relaxation.

It is further resolved that we express our indebtedness to the Brasilian scientists for providing an extensive sampling of national foods and drinks with their ever present hospitality.

We express thanks to the clerical staff for efficient and untiring efforts in our behalf, and to the photographer for his records of serious and festive events."

The Secretary reported that the Society would hold its next international congress in Canada in the summer of 1958, to consist of the Fourth International Biometric Conference and a Symposium with IUBS support on some phase of biometrical methods in genetics. Their time and place would be synchronized with the Xth International Genetic Congress, which was also scheduled for Canada in the same

summer. Following announcements by Dr. Fraga, the Symposium was declared adjourned by President Cochran. In the afternoon, those remaining after the Symposium were taken for a tour of a dairy and coffee farm near Campinas, and on the following day to Santos, the port for São Paulo on the Atlantic, both delightful excursions.

SCIENTIFIC PROGRAM

(The program of the Symposium was arranged by an Organizing Committee consisting of W. G. Cochran (Chairman), C. I. Bliss, F. G. Brieger, F. J. Crow, D. J. Finney, C. G. Fraga, J. A. Rigney and P. V. Sukhatme.

July 4. 10 a.m. Presidential address by W. G. Cochran—The 1954 trial of the poliomyelitis vaccine in the United States. (see page 527).

3 p.m. Biometrical Genetics. Chairman: F. G. Brieger. Sir Ronald Fisher—The contribution of biometry to plant breeding. E. R. Dempster—Genetic models in relation to animal breeding. F. G. Brieger—Behavior of autogamic populations and heterotic genes. H. Kalmus—Some genetical consequences of cyclomorphosis.

July 5. 9 a.m. Experimental Designs for Perennial Crops. Chairman: W. G. Cochran. S. C. Pearce—The specific problems of experimental design and technique in perennial crops. E. Amaral—The estimation of missing plots in perennial crops. A. Conagin and C. G. Fraga—Design and analysis of coffee experiments. F. Pimentel Gomes—Methods of describing crop response to fertilizers in perennial crops.

2 p.m. Experimental Designs. Chairman: G. Darmois. G. M. Cox—Recent advances in experimental designs with particular reference to estimating responses to rates of application. W. J. Youden—Design of experiments in the physical sciences.

July 6. 10 a.m. Panel Discussion on Experimental Designs for Perennial Crops. Chairman: W. L. Stevens.

July 7. 9 a.m. Statistics Applied to Animal Feeding Experiments. Chairman: B. B. Day. P. G. Homeyer—Technique and sources of variation in animal feeding experiments. G. L. da Rocha—Grazing experiments in the state of São Paulo. G. O. Mott—The grazing trial for measuring the output of pastures. A. Linder—On a particular kind of grazing experiment.

1:45 p.m. Sampling Techniques. Chairman: A. M. Flores. M. H. Hansen and J. Steinberg—Control of errors in surveys. P. V. Sukhatme—Sampling techniques for estimating the catch of sea fish. J. Nieto de Pascual—National morbidity survey in Mexico. W. L. Stevens and S. Schattan—The sampling of coffee for forecasting harvests. E. Cansado—Sampling without replacement from finite populations.

July 8. 1:30 p.m. Bioassay. Chairman: W. J. Youden. C. I. Bliss—Confidence limits for measuring the precision of bioassays. D. J. Finney—Cross-over and single-subject design for 4-point assays. O. G. Bier and P. Mello Freire—Application of bioassay to complement fixation. By title, M. Masuyama and M. Hatamura—Recent advances in biometry in Japan.

4:30 p.m. Panel Discussion on Bioassay. Chairman: A. Linder.

July 9. 9 a.m. Medical Statistics. Chairman: G. Rasch. J. O. Irwin—The study of the physiological effects of hot climates. J. Manceau—Application of the covariance analysis to the comparative study of two anthelmintics. A. E. Brandt and G. H. Fletcher—Design of a clinical investigation of very high voltage sources in the radiotherapy of cancer. A. Vessereau—Utilisation de l'analyse discriminatoire pour un diagnostic medical.

REGISTRATION

The 98 participants in the Symposium represented 17 different countries and of this group 57 are members of The Biometric Society. The following list gives the participants by countries: *Brazil* (*State of São Paulo*)—A. Sousa Quieroz do Amaral, Fernando Andreasi, H. Antunes Filho, H. Vaz de Arruda, S. Correa de Arruda, Elza S. Berquo, Otto Bier, A. A. Bitancourt, F. G. Brieger, L. de Freitas Bueno, Claudio Carvalheira, Armando Conagin, Candida H. T. M. Conagin, D. Mattas Dedecca, M. S. Dias, A. D. Netto, Constantino G. Fraga, Edison Galvao, J. F. Harrington, S. B. Henriques, Warwick Estevam Kerr, C. A. Krug, R. Aguiar da Silva Leme, Walter Leser, F. F. Manzoli, R. Franco de Mello, P. Mello Freire, A. J. Teixeira Mendes, Jose Mitidieri, Antonio Morales, A. Martins Penha, F. Pimentel Gomes, A. Mendes Peixoto, Humberto Rangel, G. L. Rocha, Victoria Rossetti, Anesiades Salati, W. R. A. Schottler, M. Rocha e Silva, J. Soubihe Sobrinho, W. L. Stevens, Odette C. Toledo, G. Pinto Viegas, Mario Zaroni; *Brazil* (*Other States*)—Edilberto Amaral, G. Garcia Duarte, Americo Groszmann, Jose Grossman, Virgilio Libonati, Ruben Markus, J. M. Pompeu Memoria, R. Meirelles de Miranda, J. Soares Neves, A. Figueiredo Penteado, F. Costa Pereira, J. B. de Barros Pimentel, A. R. da Silva, G. A. Drummond, J. N. Manceau, A. Garcia de Miranda Neto, Erik Smith, Estavam Strauss. *Argentina*—M. Guibourdenche de Cabezas. *Bolivia*—J. H. Jimenez. *Chili*—Enrique Cansado. *Colombia*—B. Romero Rojas. *Costa Rica*—Mario Gutierrez. *Denmark*—G. Rasch. *El Salvador*—Floyd R. Olive. *France*—G. Darmois, A. Vessereau, P. E. Vincent. *Great Britain*—D. J. Finney, R. A. Fisher, J. O. Irwin, H. Kalmus, S. C. Pearce. *India*—C. R. Rao. *Italy*—P. V. Sukhatme. *Japan*—T. Kitagawa. *Mexico*—Ana Maria Flores. J. Nieto de Pascual. *Portugal*—Flavio Resende. *Switzerland*—A. Linder. *United States of America*—C. A. Bicking, C. I. Bliss, A. E. Brandt, W. G. Cochran, G. M. Cox, B. B. Day, E. R. Dempster, Th. Dobzansky, M. H. Hansen, P. G. Homeyer, E. Lukacs, G. O. Mott, W. R. Pabst, W. J. Youden.

BRIEF OF PRESIDENTIAL ADDRESS:

THE 1954 TRIAL OF THE POLIOMYELITIS VACCINE IN THE UNITED STATES

WILLIAM G. COCHRAN

This trial represents an important application of biometrical principles in the struggle against disease. The experimental subjects were children in the first three grades or classes of school, of ages about 6–9 years. In terms of numbers of subjects the experiment may be the largest that has ever been conducted.

MAJOR DIFFICULTIES IN THE CONDUCT OF A TRIAL

(1) Poliomyelitis is a relatively rare disease. From past experience, the rate of paralytic polio in the study areas might be anticipated to be about 30 cases per 100,000 children aged 6–9 years. Given this attack rate, table I shows the probability of obtaining a statistically significant result (5% level) for various numbers of children and for various degrees of *true* effectiveness of the vaccine. With a vaccine that actually was 50% effective, about half a million children would be needed to make the risk of an inconclusive result small. Table II shows

TABLE I

Probability of obtaining a significant result (5% level)

No. of children in trial	True effectiveness of vaccine		
	50%	70%	90%
200,000	0.59	0.91	>0.99
400,000	0.88	>0.99	>0.99
600,000	0.97	>0.99	>0.99

TABLE II

Confidence limits for the true effectiveness

No. of children in trial	Observed effectiveness of vaccine		
	50%	70%	90%
200,000	27%—73%	34%—88%	68%—98%
400,000	29%—69%	48%—83%	77%—96%
600,000	27%—66%	53%—81%	80%—95%

the 95% confidence limits that would be obtained for the true effectiveness, if the observed effectiveness in the trial turned out to be 50%, or 70%, or 90%. Even with 600,000 children the true effectiveness can be none too well determined, except for a vaccine with an effectiveness up in the 90% range.

(2) The disease is difficult to diagnose: even in the paralytic form mistakes can be made. Some of the indefiniteness can be removed by adopting stringent criteria for the definition of a case. However, this device, if carried too far, may defeat its own ends by reducing the "accepted" cases to a very small number.

(3) The vaccination itself required 3 injections, the second given one week and the third 5 weeks after the first.

(4) The experiment subjects were children. Would parents give permission? Would physicians, health officers and medical societies give and encourage cooperation?

(5) Some biometricians have learned from bitter experience to take a pessimistic view of the prospects of success of *any* large trial with human subjects. Procedures that are essential for valid comparisons are apt to be cast aside as administratively impractical: instructions issued from a central office may be misread, misinterpreted or simply changed by persons a long way off; incomplete record forms and missing data flourish, and so on.

THE PLAN OF THE STUDY

The National Foundation for Infantile Paralysis invited the states individually to participate in the trial. If a state agreed, the vaccine was tested in all schools in certain counties within the state that had been selected by the Foundation. In order that the evaluation of the vaccine should be independent of the Foundation, the operation of the trial and the analysis of results were placed under the direction of Dr. Thomas Francis, with headquarters at the University of Michigan.

The plan announced by the Foundation was that the second-grade children in a participating school would receive the vaccine, while first and third grade children would remain unvaccinated to serve as controls.

This plan is subject to a number of potential biases. It requires the assumption that the attack rate among second-grade children is the same as the average attack rate amongst first and third grade children. Secondly, not all parents of second-grade children would allow their children to be vaccinated. Actually, 69% of them gave permission. Thus the plan compares a selected 69% of the second-grade children with the other two grades. There are epidemiological grounds for

arguing that this selection biases the results against the vaccine. Further, in any suspected case of the disease, it would be easy to discover whether a child had been vaccinated. This fact could create an unintentional bias in diagnosis by the local physician and could affect the completeness of reporting, as well as the precautions taken by parents for their children in the event of an epidemic.

It might be argued that the cumulative effect of these sources of biases was bound to be small and that results could not be seriously distorted if the vaccine was potent. But this assertion cannot be proved and with this method there must remain an element of doubt.

This plan was followed in 33 states, with 222,000 second-grade children vaccinated and 725,000 controls from the first and third grades.

A number of states adopted a different plan. Participating children in the 3 grades were divided at random into two groups. One group received three shots of the vaccine: the other received three shots of an inert fluid made up to have the same appearance as the shots of vaccine. The two treatments were distinguished by code numbers accessible only to those in charge of the study.

This plan raised more administrative difficulties than the first plan, but was free from the sources of biases that have been mentioned with respect to the first plan. All diagnoses, reporting and classification of cases, and all except the final stages of the analysis were done in ignorance of whether the child had received vaccine or placebo.

This plan was adopted in 11 states. Each treatment (vaccine or placebo) was represented by some 201,000 children. It is highly encouraging to biometricians that state officers and epidemiologists in these states expressed their preference for this plan, despite its many difficulties of execution.

Space permits mention of only a few aspects of the operation of the experiment. Collection of data was a formidable task, involving large numbers of letters, telegrams, telephone calls, regional and local conferences and special visits by members of the evaluation team to local areas. These efforts produced a high degree of completeness: missing data were of negligible importance.

Diagnoses were obtained in the following manner. When a suspected case appeared, a clinical history, including spinal fluid examination and blood and stool specimens, was made by the local physician on a standard form. A muscle examination was conducted by a physical therapist 10–20 days after onset, and a further examination 50–70 days after onset: each muscle report was reviewed by a local physician experienced in the clinical aspects of polio.

On the basis of these local records, a team of experts recruited by

the evaluation center at Michigan classified each case into one of the categories: (1) not polio (2) suspect (3) non-paralytic polio and (4) paralytic polio. The paralytic cases were further classified as to type and severity of paralysis. All these diagnoses were made by criteria that had been thrashed out and written down in advance by the team.

For record keeping and statistical analysis at the evaluation center itself, a small team of persons familiar with the handling and processing of large masses of data was obtained on leave of absence from the Bureau of the Census.

SOME RESULTS

Results were analysed and presented separately for the two plans. Areas covered by the original plan were called *observed* areas, while those that participated in the second plan were called *placebo* areas.

Table III shows the numbers of cases and the case rates per 100,000 children in the two areas. Incidentally, the paralytic case rates among non-vaccinated children were 43 in the placebo area and 44 in the

TABLE III

Cases and case rates per 100,000 children

Areas	No. of children in study	Polio cases			
		Paralytic		Non-paralytic	
		No.	Rate	No.	Rate
Placebo					
Vaccinated	200,745	33	16	24	12
Placebo	201,229	115	57	27	13
Not inoculated	338,778	121	36	36	11
Observed					
Vaccinated	221,998	38	17	18	8
Controls	725,173	330	46	61	8
2nd Grade not inoculated	123,605	43	35	11	9

observed area. Both rates were substantially above the anticipated rate of 30 which I used in discussing the needed sample size, so that the study had good fortune in not taking place during a year of unduly low incidence. The cases included in the results were all those that occurred between two weeks after the third injection and December 31, 1954.

In the placebo areas, paralytic case rates were 16 for vaccinated children and 57 for unvaccinated children. This gives an estimated effectiveness of 72%. In the observed areas the corresponding rates were 17 and 46, with an indicated effectiveness of 64%.

For non-paralytic cases, the rates were practically the same in vaccinated and control groups in both the placebo and observed areas. Although this result is somewhat unexpected, at least to a layman, it need not give concern from a public health point of view, since non-paralytic polio is not a major hazard like the paralytic form of the disease.

Table III also carries two lines marked "Not inoculated." In placebo areas this line refers to children in all three grades whose parents did not give permission to participate, plus a small number of children who received only one or two shots of placebo. In observed areas this group comprizes second-grade children whose parents did not request participation. In both areas the "not inoculated" group showed lower paralytic rates than the corresponding controls (36 against 57 and 35 against 46).

A difference in this direction had been anticipated on epidemiological grounds. Children of parents who withheld permission might be expected to be of a somewhat lower economic level than participating children, and to have acquired a greater degree of natural protection against polio through a previous subclinical attack of the disease. This type of selective bias has no effect on the results in the placebo areas, in which the comparison between vaccine and placebo was made entirely from participating children. In the observed areas, the bias would tend to reduce the apparent effectiveness of the vaccine. The fact that the vaccine showed lower effectiveness in the observed than in the placebo areas (64% against 72%) is in line with this explanation. A special sample survey that was made of participating and non-participating parents also tended to confirm the presence of a difference in economic level.

Table IV shows the estimated effectiveness of the vaccine as obtained from two more stringent criteria of classification. The main points to note are that the more severe criteria bring about some increase in the estimated effectiveness, and that the effectiveness figures run consistently about 10% lower in the observed than in the placebo areas.

The problem of making tests of significance and constructing confidence limits requires some consideration. One approach is to assume that the number of cases under a specific treatment in a school will follow a Poisson distribution. The total number of cases over all schools will then also follow a Poisson distribution, and the tests and

TABLE IV

Results given by more severe diagnostic criteria

Diagnosis	Placebo areas			95% limits for effectiveness
	No. of cases		Estimated effectiveness %	
	Vac.	Control		
Paralytic	33	115	72	57—81
Lab. confirmed	10	68	85	71—93
Positive virus obtained	15	70	80	62—89
	Observed areas			
Paralytic	38	330	62	47—74
Lab. confirmed	16	198	74	56—86
Positive virus obtained	20	210	69	50—82

limits can be constructed from Poisson theory. A more conservative approach, which avoids the Poisson assumption, is to regard the county as the basic sampling unit. The tests and limits are made by "continuous variable" theory, using the interaction with counties as the measure of error.

By either approach there is no doubt of the statistical significance of the beneficial effect of vaccine on paralytic cases. Confidence limits obtained by the Poisson approach appear in table IV, and serve to indicate the realm of uncertainty in our information as to the real effectiveness of the vaccine. The corresponding limits as obtained from the continuous variable approach would be somewhat wider.

Much credit is due to all who cooperated in this trial, and particularly to Dr. Francis and his staff, for the high standards maintained throughout the operation, despite the huge numbers of children to be processed. Among the many factors that contributed to give a fully valid comparison in the placebo areas, some of the most important were: (1) Randomization of children between vaccine and placebo (2) Keeping those concerned with case finding, diagnosis and classification in ignorance as to the treatment given to any child (3) Adoption of detailed criteria for the final diagnosis and classification and (4) Willingness to take endless pains to secure completeness and uniformity in reporting.

The question of the safety of the vaccine when given to such large numbers was of great concern. Special reports on all deaths of children,

from whatever cause, records of unusual reactions following shots, and studies of absenteeism from schools following shots were made. None of these indicated any basis for apprehension about the safety of the vaccine in this trial.

No discussion has been given here of a large volume of laboratory work designed to test the lots of vaccine, to study the rises in antibody levels following vaccination and to attempt to identify the virus from any case.

The Summary Report issued by the Vaccine Evaluation Center, University of Michigan, from which the data presented here were taken, should be consulted for a much more adequate account of the trial.

ABSTRACTS OF PAPERS

International Biometric Symposium, Instituto de Educacao Carlos Gomes, Campinas, Brazil, July 4–9, 1955

343 R. A. FISHER. **The Contribution of Biometry to Plant Breeding.**

The lecturer listed these contributions under the headings of

(1) Experimental Design
(2) Biometrical Genetics, in the sense of K. Mather
(3) Biometrical applications to classical Mendelian genetics.

He emphasised that the art of plant improvement needed in addition to genetical knowledge, the role of which will doubtless increase as greater refinement and penetration is attempted, a basic familiarity with agricultural science, and especially with the art of carrying out field trials with accuracy. Indeed the greater part of the money value due to plant improvement to date must be ascribed to improvement in field plot techniques.

After reviewing some of the concepts of biometrical genetics in the analysis of variance of the metrical values in a plant population, the lecturer turned to the increased complexity introduced into classical genetics by the study of polysomic inheritance, and the extended use in this field of complex analyses depending on observed frequencies.

344 EVERETT R. DEMPSTER. **Genetic Models in Relation to Animal Breeding.**

As gains from selective breeding diminish, attention is necessarily shifted from methods for obtaining the most improvement in a single generation to methods for achieving maximum gains over a span of many generations. On the basis of simplified assumptions, the ultimate gain from indefinitely repeated mass selections would lie between $2Nz\sigma h^2$ and $2Nz\sigma\left[1 + \frac{2n-1}{2n}\right]$, where N and n are the numbers in the population and of parents respectively, z is the height of the ordinate separating selected parents from the remainder of the population, and h^2 is the heritability. This is maximized when half, or very slightly more than half, of the population is used as parents in each generation,

and hence when selection intensity is relatively low. In moderate to large sized populations, however, this formulation would apply only to those loci where the differential effects of alleles are small to exceedingly small respectively. This demonstrates that quantitative predictions require, among other information, a knowledge of the distribution of the magnitudes of differential allelic effects at different loci. Even a qualitative conclusion—that intense selection will reduce ultimate gains—would be justified only if (as may be true) other deviations from the simplified model, such as non-additive genetic variance, linkage, and negative genetic correlation between natural fitness and characters selected for, also tend to produce a similar relationship.

It is clear that useful predictions regarding long term gains require much more knowledge than is currently available regarding variation in breeding populations. The difficulties of obtaining such information are so great that no reasonable method of attack can be neglected. One method involves deductions from what is known or may be reasonably assumed in regard to mutation rates, natural selection, and Mendelian inheritance. Such deductions have been made on the basis of simple assumptions, but only recently has much attention been paid to selection pressures variable in space and time. It is shown by an example that, under some circumstances, alleles otherwise subject to elimination by natural selection or drift could be retained in a population indefinitely if the selection pressures of given average values, were variable in time instead of steady.

345 F. G. BRIEGER. **Behavior of Autogamic Populations and Heterotic Genes.**

The requirements of applied genetics are undergoing a very significant change. In large regions of the world, mainly outside Europe and at least parts of the USA, the climatic, edaphic and economical requirements are extremely diversified and in excess to the number of people engaged in breeding work. Thus only such a degree of homogeneity seems to be desirable which is still compatible with a sufficient amount of plasticity, in order that the improved material may serve over larger areas. To achieve this, a shift of methods is necessary, giving preference to breeding methods which deal not with well defined pedigree lines, but with populations. In order to plan efficient work in population breeding, it is necessary to work out models of the genetic constitution of populations, taking into consideration the effects of recurrent mutation and selective activities, which may be applied to both panmictic and autogamous populations and to all intermediate cases which may occur.

These formulae should give the frequencies for the three basic genotypes (AA, AA', and $A'A'$).

Using a basic approach given by R. A. Fisher, we may start from equal numbers of individuals of the three genotypes, and determine how many individuals or gametes of each are left after all processes of selection have played their part. If these numbers be a, b and c, we may determine the coefficients of survival of the two homozygotes with reference to the heterozygotes by dividing by the remaining frequency b of these heterozygotes. Thus the survival value of heterozygotes is unity, and we obtain two coefficients only: $R_A = a/b$ for the homozygotes AA and $R_{A'} = c/b$, for the homozygotes $A'A'$, which may have any value from zero (complete elimination of homozygotes) to infinite (complete elimination of heterozygotes). We may then determine the triple proportions for a manofactorial population at equilibrium and for both panmictic and autogamous populations, with reference to the three most important types of mutant genes: recessive subviables, neutrals, and heterotics:

	Panmictic				
	AA		AA'		$A'A'$
Neutrals	v^2	:	$2uv$	:	u^2
Rec. subviables:	1	:	$2\sqrt{\frac{\mu}{1-R_{A'}}}$	:	$\frac{\mu}{1-R_{A'}}$
Heterotics:	$(1-R_{A'})^2$	:	$2(1-R_A)\cdot(1-R_{A'})$	:	$(1-R_A)^2$

	Autogamous				
	AA		AA'		$A'A'$
Neutrals	v	:	$8uv$	:	u
Rec. subviables:	1	:	$4u$	:	$\frac{u}{1-R_{A'}}$
Heterotics:	$(1-2R_{A'})$	:	$4(1-2R_A)\cdot(1-2R_{A'})$	:	$(1-2R_A)$

In panmictic populations all three neutral or modifier genetypes are of the same order, all being sterms of second order. In autogamous populations however the frequencies of heterozygotes are proportionally smaller than those of homozygotes, and the population will actually consist of a mixture of "pure lines". The frequency of subviable

homozygotes is small, but equal in both types of populations. The maximum frequency of heterozygotes in autogamous populations is very small and equal to four times the mutation rate only, since in each generation there appear by mutation 2μ heterozygotes, while at the same time half of all heterozygotes are lost by segregation. Heterotic loci must be very rare or even practically non-existent in self-fertilized populations, since only such loci behave as heterotics which contain two alleles, with homozygotes of a viability less than half of heterozygotes (R_A and R_A' both smaller than 0.5).

We may also use these formulae to explain the situation of special cases, such as that of heterosis in panmictic species. Thus it can be shown easily that about 1.000 loci of subviable recessives give results comparable to about 50 heterotic loci.

We may obtain also models for multifactorial segregations, by calculating the terms of the respective trinomials, with an exponent n equal to the number of loci involved. The frequencies of special cases of gene interaction may be calculated by uniting the respective frequencies of classes before termination of the calculation of the polynomial terms.

Finally we may easily determine the loss of a population caused by selection. The above formulae refer to the situation at the beginning of a generation, and by multiplying the frequencies of homozygotes by total or partial survival values and taking the differences from the original frequencies, the percentage loss may be determined.

346 S. C. PEARCE. **The Specific Problems of Experimental Design and Technique in Perennial Crops.**

Problems are considered in three classes, (1) those in scientific approach arising from the small number of experiments possible within a limited period, (2) mathematical problems in design and interpretation, and (3) those concerning the variability and measurement of plants.

Experimentation with perennial plants is laborious, and survey methods difficult of application on account of the many factors to be disentangled. A further possibility is to use sequential methods to investigate the opinions of those with the experience to give a useful judgement on a specific question, but it is still essential to [make the best use of the limited number of trials that can be done. This involves] see[ing] each experiment against the corpus of existing knowledge, [and not merely accumulating facts which cannot be explained, as is often done profitably with annuals.] The experimenter should [in fact, think about the subject, and then] use his experiment to confirm or

refute his chain of thought, [and not merely to establish phenomena in isolation. If this is the aim,] it is not enough to measure only crop [or whatever is under study;] records are required of anything that may lead to understanding of the cropping situation.

Three problems in the second group were mentioned: (a) The evolution of row and column designs, for these often enable outside trees as well as inside ones to be used in an experiment, and also permit the addition of treatments to a trial in randomized blocks. (b) The consequences of using parts of an organism, e.g., the branches of a tree, as the plots of an experiment, bearing in mind the possibility of the treatments applied to one plot affecting also the other plots of the block. (c) The combining of several years' results, a problem that awaits satisfactory progress, unlike the other two in which useful advantages have been made.

In the third group, study needs to be made of the relative importance of the various sources of variation to avoid waste of effort in controlling those of less effect. Secondly, and almost certainly the dominating problem still to be solved, is the measurement of important characters in the living tree; only when the few characters that can now be measured are supplemented will it be possible to understand how treatments have their effect, and thus to make best use of the experiments possible.

347 A. CONAGIN AND C. G. FRAGA. **Design and Analysis of Coffee Experiments.**

An outline is given of the designs and procedures of statistical analysis in coffee experiments at the Instituto Agronomico de Campinas. Four groups of experiments were considered: 1) fertilizer tests; 2) varietal trials; 3) progeny tests; and 4) miscellaneous experiments.

Under 1) the writers discussed old experiments with systematic layouts and more recent ones with randomized designs. A factorial experiment supplied conclusive results within a few years. The problem of changing some of the treatments arose in one of the experiments and the solution proposed by the writers was discussed.

Under 2) the treatment of data used by W. L. Stevens (1949) in the analysis of a systematic experiment comparing coffee varieties was commented in detail.

The evolution in designs used for progeny comparisons was discussed under 3). In early tests progenies were compared on the basis of rows of 20 plants without replications; later the designs were changed to replicated plots of 4 plants per plot, and more recently to plots of a single plant with a higher number of replications.

The results of spacing trials and of tests comparing several methods of planting coffee were discussed under 4). The plot size for coffee experiment was also discussed, based on individual yield records from a planting of the Bourbon variety.

348 F. PIMENTEL GOMES. **Methods of Describing Crop Response to Fertilizers in Perennial Crops.**

The author discusses briefly the advantages and inconveniences of fitting polynomials and Mitscherlich's equation to data corresponding to graded levels of fertilizers, the conditions which the data are supposed to fulfill in order for the fitting of Mitscherlich's law to be possible, spacing of levels to increase the probability of a good fitting, and further advisable procedure to ensure sound experimentation having in view the fitting of Mitscherlich's law.

For the special case of sugar cane, one may try to describe crop response to fertilization in a whole cycle or in each harvest separately. Factorial experiments are usually to be preferred, but in some instances just one nutrient is varied in the experiment. Such was the case in an experiment carried out in the Usina Monte Alegre (Piracicaba), by E. M. Cardoso, with the levels 0, 10, 20, 30, 40 and 50 *kg/ha* of K_2O.

Mitscherlich's equation was $y = 98.2\ [1 - 10^{-2.03(x+0.416)}]$ t./*ha*. The most profitable level of fertilization was $x^* = 58$ *kg./ha.* of K_2O. For data obtained by Strauss, in Pernambuco, in twenty-three $3 \times 3 \times 3$ factorial NPK experiments with sugar cane, the equations and most profitable levels were for Phosphorus: $y = 73.32\ [1 - 10^{-0.918(x+0.490)}]$ *t./ha.*, $x^* = 82$ *kg.* of P_2O_5 per hectare; for Nitrogen: $y = 71.97\ [1 - 10^{-0.366(x+1.789)}]$ *t./ha.*, $x^* = -39$ *kg.* of N per hectare; for Potash: $y = 65.49\ [1 - 10^{-0.984(x+0.871)}]$ *t./ha.*, $x^* = 68$ *kg.* of K_2O per hectare.

It is usual in Brazil to plant sugar cane in such a way that it has a 3 1/2 year cycle, with three harvests. The first ratoon yields around 70%, and the second around 50% of the first harvest, which is, therefore, the most important one. So, a way to solve the problem of determining the most profitable level of fertilization for each crop of the sugar cane cycle could be to obtain firstly the most profitable level for the first harvest and then, taking this for granted, study separately how much fertilizer should be used in the first ratoon. Again, assuming that the most profitable levels for the first and second harvests are used, we should try to find out what is the best level for the second ratoon. However, since in most cases fertilizers are used only in the first harvest, another way to describe sugar cane response to fertilizers would be to take together the first harvest and the two ratoons. This was done in

the analysis of a 2 × 4 × 3 factorial experiment with NPK, carried out at Usina Itaiquara by the Sugar Cane Department of the Instituto Agronomico.

The levels of P_2O_5 were 0, 60, 90 and 120 *kg./ha.* Equation obtained was $y = 316.50\ [1 - 10^{-0.859(x+1.23)}]$ *t./ha.* and the most profitable level was $x^* = 107$ *kg.* of P_2O_5 per hectare.

A further way of analysing factorial experiments with fertilizers, either for annual crops or perennials would be the fitting of a Taylor series with several independent variables, following the sequential methods of Box and Wilson. However, the author thinks that such a way is not advisable because: 1) sequential methods are too slow in agricultural research; 2) the asymptotic regression given by Mitscherlich's equation seems to be suitable in most cases, which is not true with respect to polynomials; 3) approximating polynomials obtained from Taylor's series are good only to describe local properties of a curve, and this is not enough when dealing with prices, which, when changing, shift our attention to other portions of the curve. In agriculture a similar shifting may be caused also by the introduction of new varieties or by other means of increasing the yield.

349 GERTRUDE M. COX. **Some Recent Advances in Experimental Designs, with Particular Reference to Estimating Response Surfaces.**

A brief survey of the basic, or classic, designs used in experimentation was given. The second section of the paper presented variation in these basic designs along with designs currently being developed. Those discussed were (1) balanced groups with covariance (2) change-over trials (3) doubly balanced incomplete block designs (4) partially balanced incomplete block designs with two associates (5) chain block designs and (6) paired comparisons.

The third and major portion of the paper dealt with the designs being used to secure an estimate of the optimal point and to explore the nature of the response surface in the vicinity of this optimum. Examples were given illustrating the use of composite and rotatable designs in actual experiments.

350 W. J. YOUDEN. (National Bureau of Standards, Washington, D. C.). **Design of Experiments in the Physical Sciences.**

Experiments in the physical sciences reveal certain features not usually exhibited by classical designs used in agricultural field trials.

(1) Block size is often uniquely determined by the apparatus, or physical equipment, or material investigated. (2) Block effects are almost always of interest and are sometimes of primary interest. Blocks may be instruments, or positions in the equipment, and these are enduring entities. (3) The relatively high precision of physical measurements means that little replication is needed and designs involving considerable replication are unacceptable. (4) Sometimes it is highly desirable not to specify all the treatments in advance because the experimenter desires to obtain some results, say at a few temperatures, before specifying the other levels. This leads to the use of designs which may be incorporated in a larger design when the first results are in hand. (5) Measurements are usually obtained in a time sequence and provision for instrumental or environmental drift is frequently necessary. (6) Finally the experimental situation is sometimes of a unique nature. For example, the environmental temperature instead of being held constant may, for good reason, be made to rise steadily. Special designs allow more precisely for the temperature effect than simple block arrangements.

351 GERALDO LEME DA ROCHA. **Grazing Experiments in the State of São Paulo.**

Methods of pasture management are of great importance for the dairy and beef cattle industry in the state of São Paulo. A number of experiments are being conducted in various parts of the state by the author and collaborators. An outline of the experiments concerning rotational grazing is being given here.

I. *Experiment in the Paraiba River Valley Region*

Four species of grasses, (Gordura grass—*Melinis Minutiflora Beauv.*; Jesuita grass—*Axonopus compressus* Beauv.; Colonião de Tanganika grass—*Panicum maximum* Jacq.; Sempre Verde grass—*Panicum maximum* var. *gongyloides* Jacq.), were compared in grazing tests using yearling dairy cattle. Thirty-two paddocks of 5,000 sq. m. each were utilized. Although no replication was made, the experiment was divided in 4 blocks of 8 paddocks each. Within the blocks the treatments were randomized. Each species of grass was planted with and without fertilizers.

Three animals in each of the 8 paddocks in block *I* started grazing on the same day and were rotated according to the condition of the sward and liveweight variations. Weighing was done once a week. Data obtained after 9 months showed no correspondence between

liveweight variation and subjective judgment of the sward. From then on the animals were left grazing in each paddock for 8—14 days and were then moved over to the next paddock (any surplus grass was grazed by "followers"). Botanical analyses were made twice a year (December-January and July-August).

II. *Experiment near Ribeirão Prêto (red soil)*

The same layout of the experiment described above was followed. One hectare paddocks of the following species of grasses were compared: Colonião grass—*Panicum maximum* Jacq.; Makari-Kari grass—*P. coloratum* L.; Gordura grass, and Jaraguá grass—*Hyparrhenia rufa* (Nees) Stapf. Grazing in this case was done with groups of 10 yearling beef cattle.

III. *Nova Odessa Experiment*

The experiment in this area aimed at finding out the best management for gordura grass swards on the basis of *use* and *rest*. Groups of 6 animals were put to graze in 5,000 sq. m. paddocks, according to the following schedule:

Treatment A—4 days use—20 days rest
B—6 days use—30 days rest
C—8 days use—40 days rest
D—continuous grazing

Botanical analyses were carried out as described before.

IV. *Collina Stud Farm Experiment*

This experiment was similar to III, but had an improved design. Sixteen 5,000 sq. m. paddocks were available. Four replications of each of 4 treatments were compared:

Treatment A—4 days use—12 days rest
B—6 days use—18 days rest
C—8 days use—24 days rest
D—continuous grazing

Grazing was done with mares. During the resting period the animals grazed on the common swards of the farm. Botanical analyses were made as before.

V. Another experiment is being carried out in the Paraiba River Valley Region. Kikuiu grass—*Pennisetum clandestinum* Hochst ex Chiov.,

and Rhodes grass—*Chloris gayana* Kunth, are being tested in two different ways: rotational grazing versus strip grazing. This comparison is based on experiments designed by Holmes and carried out at the Dairy Hannah Institute, Ayr, Scotland.

352 G. O. MOTT. (Purdue University, Lafayette, Indiana). **The Grazing Trial for Measuring the Output of Pasture.**

The objective of the grazing trial is to measure the quality of herbage produced by a pasture and the yield of animal product per unit area. It yields information useful to both the Agronomist and Animal Husbandman in that the output per animal is an indication of the quality of forage, the carrying capacity in terms of animal days is a reliable index of the herbage production per unit area, and the livestock product per acre is an indication of both the quality and quantity. The daily output per animal is a function of the nutritive value of the forage, the rate of intake and the physiological characteristics of the animal. The performance of the animal is also greatly affected by the grazing pressure and the opportunity for selective grazing.

The size and sources of the experimental errors associated with the grazing trial differ for the several units of measure. Both the pasture variability and that which can be attributed to the animal has to be considered. A study of the sizes of the errors for the various units of measure points to the need for at least three and preferably five or more field replications in a grazing trial. The size of the field for each replication should be sufficient to supply at least two animals with adequate herbage.

The source of bias most commonly encountered in the grazing trial is the failure of the investigator to estimate carrying capacity at the optimum. If the pasture is overgrazed, the number of animal days will be overestimated. the daily performance of the animal will be less than that expected at the optimum and the product per acre will also be underestimated. If on the other hand the pasture is not grazed to capacity, then the number of animal days will be low, the daily gains may be slightly overestimated and the product per acre will be underestimated due to the failure to utilize the forage produced.

Simple designs are usually indicated for the grazing trial due to the limited number of treatments and replications involved. Techniques used to reduce the errors due to the animal in the feeding trial are also useful in the grazing trials for variates such as previous performance and initial weights.

353 MORRIS H. HANSEN AND JOSEPH STEINBERG. (U. S. Bureau of the Census, Washington, D. C.). **Control of Errors in Surveys.**

In the evaluation of the Current Population Survey—(a monthly population survey made by the U.S. Bureau of the Census to estimate labor-force characteristics and other population data)—control of errors has been sought through the traditional devices of selection, training and supervision of personnel. Efforts to increase the objectivity of the measurement of error in the enumeration and interview processes are pursued by a formal quality control procedure based on re-interviews.

During a twelve-month period the recheck of coverage indicated minimal errors by a preponderance of the enumerators (85% of enumerators had zero error rate) and concentration of errors within a small portion of the enumerators (5% of enumerators were the source of approximately three-quarters of the errors); however, the recheck of information obtained during the interview indicates that approximately 60% of the enumerators have varying amounts of differences between the original and recheck results. Various possible causes of discrepancies in content were examined. The respondent appears to contribute as much or more heavily to the differences than does the interviewer. The sources of error are difficult to identify and consequently the check results are difficult to interpret in controlling the work of individual interviewers. Examination of net differences between original interviews and reinterviews in terms of estimated standard errors indicates that for the most part the interviewing in the Current Population Survey can be considered to be under control.

Experimentation is continuing on various sources of error and the control of observational errors.

354 P. V. SUKHATME, (FAO) and V. G. PANSE, (ICAR). **Sampling Technique for Estimating the Catch of Sea Fish.**

The paper describes a sampling method developed in India for estimating the monthly catch brought to the coast by fishing boats. It is divided into two parts—the first dealing with the sampling procedure for estimating the daily catch at selected sections of the coast and the second dealing with the choice of optimum first-stage unit for sampling and the number of such units for estimating the monthly catch for the entire coast with given precision.

The sampling procedure for estimating the daily catch is based on the study of hourly landings at a number of selected sections along the

coast, and it is concluded that a systematic selection involving two visits of three hours (or two, in case the journey time to the coast is shorter) during a day is both a practical and efficient scheme of sampling at selected sections. The section of the coast found most suitable for use as sampling unit for observation is a landing centre. The paper then gives the result of a study to determine the optimum number of days for which a selected centre should be observed in succession, based on the data collected at 61 landing centres for two months and concludes that a centre × day is the optimum unit of observation. The number of centres to be selected daily for estimating the monthly catch for the entire coast with 5 per cent error is placed between 15 and 22.

Finally, a brief description is given of the surveys conducted in India in the course of which the above technique was developed. The present surveys cover a coast of nearly 500 miles in length. It is proposed to extend the surveys to cover the entire coast of India as the normal method of estimating the monthly catch of marine fish.

355 ENRIQUE CANSADO. **Sampling Without Replacement from Finite Populations.**

Although Hansen and Hurwitz (1953), Midzuno (1950), Narain (1951), Horwitz and Thompson (1952) and Yates and Grundy (1953) made important contributions to a general theory of sampling without replacement from finite populations, the presentation of this general theory did not satisfy the standards with respect to rigour and systematicity, that are now prevalent in the expositions of other branches of Calculus of Probability and Mathematical Statistics.

In this paper this well-known theory is presented in a form which, it is hoped, will reduce some of the aforementioned deficiencies. Starting with the fundamental set of selecting probabilities $P_1(u_i)$, $P_2(u_i/u_j)$ $\cdots$, $P_n(u_i/u_{h_{n-1}}, \cdots, u_{h_1})$, which defines completely the sampling scheme considered, formulae are given for the probabilities $P_1(u_i)$, $P_2(u_i)$, $\cdots$, $P_n(u_i)$ of selecting the unit u_i at the first, second, $\cdots$, n-th draw. From these are obtained the probabilities $P(u_i)$ of inclusion of the unit u_i in a sample of size n. Formulae are also considered which give, from the fundamental set of selection probabilities, the probabilities $P_{r,s}(u_i u_j)$ of selecting the unit u_i at the r-th draw and the unit u_j at the s-th draw. From these are obtained the probabilities $P(u_i u_j)$ of including both the units u_i and u_j in a sample of size n.

On this basis it was easy to obtain the formulae for the mathematical expectations of the sums and the product-sums of the observed values in

a sample of size n. Then $\hat{T} = \sum_{r=1}^{n} \frac{x_r}{P(u_r)}$ was considered as an estimator of the population total $T = \sum_{i=1}^{N} X_i$. It was, then, readily shown that $\hat{T}$ is unbiased and formulae are obtained for the sampling variance $V(\hat{T})$ of this estimator.

The paper ends with a consideration of two unbiased estimators of this sampling variance:

$$\hat{V}_1(\hat{T}) = \sum_{i=1}^{n} \frac{1 - P(u_i)}{P^2(u_i)} x_i + \sum_{i \neq j=1}^{n} \frac{P(u_i u_j) - P(u_i)P(u_j)}{P(u_i)P(u_j)P(u_i u_j)} x_i x_j$$

given by Narain and Horwitz and Thompson, and

$$V_2(\hat{T}) = \frac{1}{2} \sum_{i \neq j=1}^{n} \frac{P(u_i)P(u_j) - P(u_i u_j)}{P(u_i u_j)} \left\{ \frac{x_i}{P(u_i)} - \frac{x_j}{P(u_j)} \right\}^2$$

given by Yates and Grundy.

356 C. I. BLISS. **Confidence Limits for Measuring the Precision of Bioassays.**

In measuring the precision of an assayed potency, confidence or fiducial limits have the advantage over the standard error of taking full account of the size of the assay and the precision of its slope. Experimentally, bioassays can be divided into two types, (1) those based upon the mean threshold dose, measured directly in each test animal, and (2) those where the size of the reaction at selected dosage levels is the dependent variable and relative potency must be inferred by converting the response back to units of dose or log-dose. For the few assays of the first type, the confidence interval of the log-relative potency M' is that for a mean difference or for the difference between two means, both well known and simple calculations. For the much larger group of assays comprising the second type, potency depends upon the ratio of two statistics. If the dosage-response curve is linear with arithmetic dosage units, potency is computed from the ratio of two slopes, one for the Standard preparation and the other for the Unknown. If instead the response plots linearly against the log-dose, the more common type, log-potency is computed from the ratio of a difference in the mean response ($\bar{y}_U - \bar{y}_S$) to a slope (b).

One of the simplest forms for computing the confidence limits for a ratio was that introduced in 1944 by Marks for balanced cross-over assays for insulin. His equation has been generalized for all assays based upon the ratio of two statistics with little loss in its inherent

simplicity. If the log-relative potency is computed as $M' = (\bar{y}_U - \bar{y}_S)/b = a/b$ and the assay is balanced so that the numerator and the denominator are independent, the confidence limits for M' can be determined as

$$X_{M'} = CM' \pm \sqrt{(C-1)(CM'^2 + v_{aa}/v_{bb})}$$

where $C = b^2/(b^2 - v_{bb}t^2)$, and the variances of a (v_{aa}) and of b (v_{bb}) have a ratio depending upon the design of the assay and is determinable in advance. Factorial log-ratio assays follow this pattern, including assays in balanced pairs and with more than one Unknown. It is also applicable to all-or-none assays and to assays based upon the ratio of two mean threshold doses.

When the numerator and denominator of M' are not known to be independent, their covariance leads to an additional term both inside and outside the radical in the above equation. The effect of this covariance in assays arranged in randomized blocks or groups, in assays needing replacements, in assays with a single test animal having a changing sensitivity, and in slope-ratio assays are considered. Illustrative numerical examples have been drawn in part from recent studies connected with the new U.S.P. XV.

357 D. J. FINNEY. (The University, Aberdeen, Scotland). **Cross-Over and Single-Subject Designs for 4-Point Assays.**

Many biological assays can be increased in precision by measuring responses to different doses successively on the same subjects. Adoption of different dose sequences for different subjects produces cross-over designs. The statistical analysis of the results of these involves some consideration of time series. In particular, the possibilities of correlation between components of residual error, of residual effects of past doses, and of autoregressive influences of one response on its successors need to be considered. This paper presented five different models that may be appropriate to bioassays.

Using these models, the analysis of 4-point parallel line assays for various assay designs (twin cross-over and Latin square types) was then discussed. Lucas's findings that the existence of residual influences need not bias these designs in respect of treatment comparisons were confirmed, and the special procedures for estimating error variances that Patterson suggested were developed. New features peculiar to bioassay, notably the presence of several independent or semi-independent validity tests and estimates of relative potency, were considered in detail.

A new class of designs for single-subject assays, possessing a property of serial balance, was described, and the statistical analysis of these was explained.

358 O. G. BRIER, M. SIQUEIRA AND P. M. FREIRE. **Application of Bioassay Methods to Complement Fixation.**

Quantitative complement fixation tests were performed with varying dilutions of sera from syphilitic patients and a constant optimal amount of cardiolipin antigen. A large, constant amount of complement was available in the fixation mixtures, and from the spectrophotometric titration of the residual hemolytic activity, the numbers of 50 per cent hemolytic units of complement bound to the various amounts of antibody were estimated.

When the dose of antibody-containing serum was expressed in the logarithmic scale, over a certain range there was a linear relation with the number of complement units fixed by any constant volume of the mixture. The adequacy of this linear relation was established in replicated tests with 4 syphilitic sera. As tested by analyses of variance, the data for each serum agreed with the corresponding regression lines within the limits of chance variation.

The reproducibility of assay results and the accuracy of potency determinations were investigated by repeating factorial 2 × 2 assays, in which one of the 2 sera being compared was a known dilution of the other. With the calculated regression lines not departing significantly from parallelism in each of 3 independent assays, the estimated potency ratios in 2 assays agreed with each other and with the true ratio. The remaining assay gave a potency ratio which significantly differed from the true value and was inconsistent with the other two estimates, as measured by chi-square. Taking this heterogeneity into account, the results of the 3 independent estimates were combined, and the average potency ratio did not significantly differ from the true ratio.

The applicability of the method to the comparison of different syphilitic sera was further examined in non-replicated tests. Regression lines were calculated for the data corresponding to 8 sera tested with 3 doses each, and a combined analysis of variance showed that the slopes of those lines did not depart from parallelism more than could be expected by chance alone.

The foregoing results indicate the possibility of assaying the antibody of syphilitic sera by complement fixation in a system of parallel straight lines.

359 J. O. IRWIN. **The Study of the Physiological Effects of Hot Climates.**

One of the physiological requirements for health is the maintenance of a practically constant body temperature. The environmental factors which affect the rate of heat loss are the temperature, humidity and rate of movement of the air and the radiation from the surroundings; but the rate of loss is largely governed by the physiological mechanisms which serve the body as thermostatic controls.

A single index of thermal environment known as Effective Temperature was designed to take account of the temperature, humidity, and rate of movement of the air and was a measure of subjective feelings of comfort. On account of defects in Effective Temperature McArdle and colleagues in London were led to construct an alternative index based on sweating rates. Using the results of nearly 1000 individual experiments they constructed an empirical nomogram from which the Predicted 4 hour Sweating Rate—*P4SR*—for any set of working conditions could be ascertained provided the environmental factors, the metabolic cost of the work and the clothing worn were known. It was desired to estimate the accuracy of the *P4SR* scale, and to assess the value of the Effective Temperature scale for grading the severity of thermal conditions in relation to human activities in the Tropics.

An experiment was carried out at Singapore to determine the effects on men naturally acclimatised to the Tropics of exposure for four hours twice weekly to varying combinations of air temperature, humidity and air movement. Combinations of air velocity, dry bulb and wet bulb temperature were designed to cover the same range as had been investigated in London. Originally a $4 \times 3 \times 2$ factorial arrangement had been suggested, but this was modified for technical reasons. There were 3 teams with 4 subjects in each—young naval ratings who volunteered from ships or shore establishments on the Far East Station. Each team had two 4-hour periods a week in the hot room. Four work-clothing combinations were tested at each exposure: Working in shorts, Working in overalls, Resting in shorts, Resting in overalls. Work consisted in step-climbing according to a certain routine. These four categories have been called "Postures" for convenience. They may be allocated to 4 subjects in 24 different ways, and one of these was assigned randomly to each of the 24 climate combinations, separate randomisations being used for each team.

In this plan all separate climate and posture comparisons were unconfounded with differences between persons and were therefore

equivalent to comparisons on the same persons. The error term for these was originally intended to be based on such climate-posture interactions as were not confounded with personal differences. It was not realised that these interactions were as important as they proved to be. Further, by what was subsequently recognized to be an error of judgment, the 12 subjects were, on the basis of a uniformity trial carried out before the main trial started, divided into four grades of sweating with three subjects in each and one member of each grade was put in each time. The arrangement of the climatic variables was thus not factorial, and it was not possible to allow for personal differences by the analysis of variance itself. The only way to correct for these was by analysis of covariance on the basis of the uniformity trial. To meet the other difficulty, the results of the trial were divided into two distinct sections, the first containing all combinations of 90° and 120° dry bulb temperature, 80° and 85° wet bulb temperature and the four air velocities, and the second all combinations of 90, 100, 120°F. dry bulb temperature and 80, 83, 85, 88°F. wet bulb temperature at the third air velocity (300 ft./min.). The analysis of covariance was carried out separately for the two sections. The combination of 90° and 120°F. dry bulb temperature with 80° and 85° wet bulb temperature at an air velocity of 300 ft./min. occurred in each.

The statistical analysis was carried out for a number of response variables—Total sweat loss, Total sweat loss per square metre of body surface, Evaporative water loss (absolute and per square metre), Final rectal temperatures, Final pulse rates (seated and standing), Comfort ratings and Efficiency ratings. Examples given refer to the variate "Total sweat loss".

Regression analysis was used to compare the "Total sweat rates" obtained from this series of experiments with the *P4SR* values obtained from the nomogram constructed by McArdle and his colleagues and with Effective Temperature. If each work-clothing combination is taken separately, the predictive accuracies for these "naturally acclimatised" naval ratings of the Effective Temperature scales and the *P4SR* nomogram are about the same, though there is a slight advantage to the latter in predicting sweat loss. However, when the results of all groups of experiments are combined, correlations with effective temperature are considerably lower than with *P4SR*, because the Effective Temperature scales make no allowance for differences in work rates. In this sense *P4SR* is a more comprehensive index. It also gives a more adequate picture of the change in stress with air movement. On the other hand, the predicted 4 hour sweat rate can only be applied within the range of climate-work-clothing combinations which cause

people to sweat. It cannot replace Effective Temperature under the more comfortable and desirable conditions of light and sedentary work with which it was designed to deal primarily, for sweating will not occur under these conditions. In this sense it is less comprehensive than Effective Temperature but it is a more accurate index of physiological effect under conditions of thermal stress.

360 J. N. MANCEAU. **Application of the Covariance Analysis to the Comparative Study of Two Anthelmintics.**

Prevention and treatment of the several helminth infestations is one of the major concerns of public health officers in the Amazon region. A test was made to compare the efficiency of Aralen a new drug, with that of Hexylresorcinol, the one commonly used in the treatment of infected persons.

A sample of 74 children was chosen at random from the Lauro Sodre professional school (Belem, State of Para), and a stool specimen was taken from each child to determine the degree of infestation (number of eggs per centigram of feces) with A. lumbricoides, Ancylostoma, and T. trichiura, before treatment.

The sample of 74 children was arranged into 37 pairs, each pair having, insofar as possible, the same degree of infestation by A. lumbricoides. One child selected at random from each pair was treated by Hexylresorcinol, and the other was treated by Aralen.

The results were subjected to the analysis of covariance. Hexylresorcinol proved to be better than Aralen in the treatment of infection with A. lumbricoides and Ancylostoma. No significant difference was observed in the treatment of infection with trichiura.

The independent variable used (degree of infestation before treatment) made it possible to attain a higher degree of accuracy in the experiment.

361 A. E. BRANDT AND GILBERT H. FLETCHER. (Biometrician, Health and Safety Laboratory, New York Operations Office, U. S. Atomic Energy Commission, and M.D., Pathologist, M.D. Anderson Hospital, The University of Texas, Houston, Texas. **Design of a Clinical Investigation of Very High Voltage Sources in the Radiotherapy of Cancer.**

The full title of this contribution should read, the design of a clinical investigation of the differences in reaction and clinical response between Cobalt-60 (1.2 Mev) therapy and 22 Mev Betatron therapy in the

treatment of cancers which are infrequently curable by conventional radiotherapy techniques. The two phases of this problem (the medical and the biometric) are defined and the paramount importance of the medical phase due to the use of human subjects is pointed out. The responsibility of the medical leader of this investigation for the medical excellence of the design as well as for the diagnosis and treatment of patients is presented. The objectives of the biometrician and the controls by which he achieves these objectives within the medical framework provided by the medical leader are given. The biometric design of that portion of the investigation relating to cancer of the cervix is presented as an example.

362 A. CHARBONNIER, B. CYFFERS, D. SCHWARTZ, A. VESSEREAU. **Application of Discriminatory Analysis to Medical Diagnostic.**

On individuals which belong to one among two or several families, measurements of several characters, or variables, have been made. The point is to find the linear functions of these variables—namely the "discriminant functions" by which we can allot each individual to its proper family with a minimum of risk of error.

Hypotheses are that, within each family, the distribution of the variables is a normal distribution with p variables, and that all these distributions have the same dispersion-matrix, but differ only regarding their centres.

In the following application of the method to the medical diagnostic, the families differ by the nature of a basic disease named "ictere". The variables are different constituents of the blood serum:

$$\text{albumin} = A \qquad \text{globulin} = \alpha_1 , \alpha_2 , \beta, \gamma, \sigma, \epsilon$$

p = total mass of proteins. $p = p - \sigma - \epsilon$.

The values of these elements, measured by the technics of electrophoresis were recorded on a group of 197 ill persons.

The study was made in several steps.

First step: Observation of evident differences between the families, concerning each variable. This empirical information can be used to predict the nature of the disease. But, in better circumstances, diagnoses were made only in 40% of cases with 8 errors (or 20% of the diagnoses).

Second step: by t test of Student-Fisher, verification of significant differences between the families, for each variable of the electrophoresis.

Third step: Choice of the best variables for the establishment of the discriminant functions. The choice was possible between x (raw variable), $\log x$, $x\ P'$ or $x\ P$. The values of t (comparison between hepatites and cancers) show that it is better to work with the variables x/P'. With six variables, plus the total mass of protides P, the computation of the coefficients of discriminant functions would have been very tedious. Only 4 variables were retained, which, regarding the values of t, were a priori the best for discrimination.

Fourth step: Computation of the discriminant functions. With 3 families, one discriminant function only is available if the centres of the families are on the same straight line. That was not the case, so it was decided to make the discrimination in two steps:

first, discrimination between "medicaux" and "chirurgicaux" (calculs + cancers) which is the most important,

then, among the "chirurgicaux", discrimination between "calculs" and "cancers".

Computations have been made with the data of the 197 diseased. Statistical tests show that, for the first function, the coefficients are significant, or almost significant, and that for the second, the coefficients are not very significant.

In order to allot *each diseased* to one of two families, a critical value is chosen. If, for a particular diseased, the value of the discriminant function is smaller than the critical value, the conclusion is: "family 1"—in the other case the conclusion is "family 2". With the best critical value, it is found that the theoretical % of errors is very high: 35%. In fact on the 197 diseased, 62 errors (or 32%) were found.

It is better to fix a priori the probability of error. With this position a segment (ab) is determined, and between the values a and b, no diagnosis is pronounced.

5% was chosen as probability of error for each of the discriminant functions separately—When the two functions are successively applied, with the possible responses:

No diagnosis
Hepatitis
Calcul or cancer
Calcul
Cancer

the probability of error is theoretically higher. The theoretical proportion of diagnoses (considering calcul or cancer as a diagnostic) is 45%; on the 197 diseased we find 41%.

Applying successively the two functions to the 197 diseased, we obtain the following results:

81 diagnoses (42%),
10 errors (5% of the total, 12% of the diagnoses).

There is a very good concordance between theoretical and observed values.

But a better confirmation of the validity of the method is the following. The discriminant functions were applied to another group of 81 diseased and the classification operated by the functions was entirely in harmony with the theory and with the results of the first group of diseased.

In conclusion, we can emphasize the following points:

1°)—This method is better than the empirical method: more responses, less errors. Its application is easy: tables which have been constructed, give very easily the values of the discriminant functions.

2°)—The method requires some precautions: it depends on the technics applied, and perhaps on the origin of the diseased.

3°)—The method has been good for the discrimination between "icteres médicaux" and "chirurgicaux", but fairly poor for the discrimination between "calculs" and "cancers". An improvement of the method can be expected, taking account of the sex, the age, of the ill-persons.

4°)—It seems that the same method would be of interest in many other cases of medical diagnostic.

363 ARTHUR LINDER. **On a Particular Kind of Grazing Experiment.**

Research has been carried out on the effect of fertilizers on pastures in the higher regions of the Grisons (Switzerland). Grazing experiments were set up to evaluate the palatability of grass on fertilized plots. One typical experiment consisted of five blocks with six plots receiving different kinds and amounts of fertilizers. A fence was drawn around the area and two cows were allowed to graze for two hours. The effective grazing times were recorded. Analysis of the results shows significantly longer grazing on plots with higher levels of fertilizers. The experiment was repeated after one year without changing the treatment of plots. Results of the two trials agreed closely.

THE BIOMETRIC SOCIETY

Italian Region. On April 22, 1955, the Italian Region held its fifth annual meeting in Pavia in a joint session with the Italian Genetic Society (Associazione Genetica Italiana). Three papers were presented. The first, by F. Brambilla and L. L. Cavalli-Sforza, on Biological variability of environmental origin, concerned a model for the non-genetic variability in a population and its statistical consequences. The method is based upon a transformation of the frequency distribution of the environmental stimuli by a function connecting the stimulus and the intensity of biological response. When this function has a maximum or minimum, the frequency distribution of the response shows some peculiar characteristics which were discussed. The second paper by A. Previtera concerned the measurement of frailty in children. Two established auxological indexes were examined statistically with a population of children of various ages and an improved index suggested. The third paper by E. Baldacci, G. Fogliani and E. Betto described the planning of experiments on the control of *Peronospora* by fungicides. The problem here lay in the fact that artificial infection with this species is not easy, so that field experiments have to be based upon natural epidemics.

British Region. The Region held its twenty-fourth meeting at the Wellcome Research Institute in London on May 4, 1955. In the first contribution, J. A. Fraser Roberts discussed the supposed difference between the sexes in variability of intelligence. That boys are more variable than girls in their scores on intelligence scales seems to arise empirically from the data. After excluding a large volume of unsuitable records, this greater variability of boys seems to be real, at least in some kinds of tests and at certain ages. The difference may be a biological phenomenon, depend on the way test scales are constructed, or be due to differences between boys and girls in education and in their social and home environments. The problem is being examined by analyzing a number of different samples. Dr. Roberts' paper was followed by a discussion opened by J. W. Craven and P. Olden on some difficulties in sequential analysis.

Switzerland. The Swiss Section of the Biometric Society met jointly with the Swiss Society of Genetics at the University of Berne on May 22, 1955. The following papers were presented: S. Rosin, Statistical problems in the evaluation of blood-group determinations; H. L. LeRoy, Mathematical statistics as an aid in the solution of problems of selection in animals; A. Kaelin, Influence of selection upon

the estimation of gene frequencies among siblings in human genetics; and A. Linder and B. Grab, A statistical study on the relation between several anthropological measures in the infant and the corresponding measurement in its parents.

Region Francaise. Lors de la reunion de la Societe Francaise de Biometrie, qui eut lieu mardi le 24 Mai a l'Ecole Normale Superieure a Paris, l'ordre du jour etait le suivant: R. Turpin et M. P. Schutzenberger, Remarques sur un probleme de consanguinite; et Dr. Charbonnier, B. Cyffers, D. Schwartz et A. Vessereau, Discrimination entre icteres medicaux et chirurgicaux a partir des resultats de l'analyse electrophoretique des proteines du serum.

WNAR. On the first day of the annual meeting of the Wester North American Region in Pasadena, California, on June 23, 1955, the program consisted of two scientific sessions and a luncheon business meeting. The morning session on Ecology, under the chairmanship of R. O. Erickson, offered the following papers: P. E. Fields, Factorial designs and the guidance of downstream migrant salmon and steelhead trout; D. G. Chapman and R. Pyke, Statistical theory of some migration population models; J. L. Baily, Jr., Variation of the *Pectin gibbus* complex; and R. F. Tate and R. L. Goen, Minimum variance unbiassed estimation for a truncated Poisson parameter. During the luncheon business meeting, following other business, the Region voted to elect future officers by mail ballot. At the afternoon session on Psychometrics, J. P. Guilford presided and the following papers were presented: H. D. Kimmel, Reliability of qualitative, categorical judgements; D. A. Grant, Statistical tests in the comparison of curves; P. R. Merrifield, Quantification of ordering behavior; J. A. Gengerelli, Methods of constellation analysis; H. H. Harman, Some observations on factor analysis; and J. W. Frick, Effect of varied interpolated stimuli upon the time-order error. At the closing section on Genetics on June 24, A. H. Sturtevant was in the chair and the following papers were read: L. A. Lider, A group of long-term, perennial and non-replicated rootstock trials; C. N. Stormont, Estimates of frequencies of B-alleles in three breeds of dairy cattle; G. E. Dickerson, Some unsolved statistical problems of importance in quantitative genetics; and H. Rubin, Axiomatization of genetics.

Japan. Members of the Biometric Society in Japan held their second spring meeting in Tokyo on April 5, 1955. Five papers were presented: M. Masuyama, Microbiological inspection of bulk material by composite sample; M. Kiyoku, On the response-surface in the interaction of temperature and dryness upon insects; T. Okuno and T. Sasaki, Factorial analysis of the adjusted treatment means obtained

by analysis of covariance; T. Yamada, On the chance distribution of quantitative characters in a population due to interplant competition; and S. Hashiguchi, Estimating the variance component for the evaluation of heritability. To insure their wider distribution, the Chapter has printed an extended summary of each paper in a 34-page Japanese booklet, in which three of the papers have been provided with English abstracts. The preceding spring session of the Japanese Chapter met jointly with the Japanese Agricultural Society.

Australasian Region. The Region held its fifth meeting at the University of Melbourne on August 22, 1955, as part of the 31st meeting of the Australian and New Zealand Association for the Advancement of Science. The following program was cosponsored by the Zoology Section of the Association: R. T. Leslie, A statistical approach to the physiological problem of thresholds; E. J. Williams, Sidelights of sampling surveys; G. S. Watson, Missing and mixed-up values in contingency tables; P. J. Claringbold, Discriminant analysis in the interpretation of semi-quantal data; and (Mrs.) G. L. Richardson and F. E. Binet, Discriminant analysis of species of the genus Murravia (Brachiopoda, Tertiary, Recent).

ENAR. The Eastern North American Region met in East Lansing, Michigan, on September 6–8, 1955, as part of the annual meeting of the American Institute of Biological Sciences. The program was arranged by a committee consisting of E. L. Green (Chairman), M. Whittinghill, T. Park and W. C. Jacob. The Local Representative was W. D. Baten.

A joint session with the Genetics Society of America and the American Society of Human Genetics on the afternoon of September 6 was titled, "Sewall Wright's Contributions to Population Genetics". Professor Wright was present and received a standing ovation by the several hundreds in attendance. The following papers were presented under the chairmanship of Paul R. David: C. C. Li, The concept of path coefficients and their impact on population genetics; J. F. Crow, Effective population number, its estimation and relevance as one factor in evolution; H. B. Glass, Some evidence for genetic drift in human populations; W. P. Spencer, Natural populations of Drosophila and the Wrightian model of evolution; and J. P. Scott, The analysis of quantitative traits in populations.

The following afternoon a joint session with the American Society of Horticultural Science was devoted to "Sampling Applications in Horticulture" with J. C. Jacob in the chair and the following papers: J. P. McCollum, Sampling tomato fruits for composition studies; W. C. Kelly, Sampling vegetative portions of vegetable plants for vitamin

analysis; W. W. Jones, Sampling citrus and avocado trees for nutritional studies and yield relationships; N. J. Shaulis, Sampling small fruit for composition and nutritional studies; and J. A. Rigney, Sampling soil for composition studies. A Biometric Society Dinner in the evening was well attended.

The closing session on September 8, held jointly with the Ecological Society of America and the American Society of Naturalists, concerned "Quantification in Population Ecology", with Thomas Park as chairman. The program consisted of the following papers: J. Neyman, Statistical models of population phenomena; D. E. Wohlschlag, Conceptual problems in the application of theoretical models to unstable fish population; and L. C. Cole, Inductive procedures in quantitative ecology.

Finances for 1954. Following the practice initiated last year, the following audited financial statements for BIOMETRICS and for the office of the Secretary-Treasurer for 1954, are listed under headings similar to those used last year.

BIOMETRICS

Income

Reserve from sale of No. 1 in Vols. 3 and 7	$ 1321.50	
Balance from 1953 books	4993.95	$ 6315.45
Member subscriptions		
ASA, 394	1576.00	
Biometric Society, 1144	3936.00	5512.00
Non-member subscriptions, 622		4352.25
Sale of back issues		
Vol. 1–5, ,864.40; Vol. 6–10, ,1195.30	2059.70	
Vol. 3, No. 1, ,217.50; Vol. 7, No. 1, ,66.00	283.50	2343.20
Sale of reprints		1124.61
Insurance overpayment (refund)		26.56
Total Income		$19674.07

Expenses

Overpayments and cancellations		$ 47.50
Supplies, miscellaneous		38.27
Farm Bureau, insurance on back issues		48.96
Postage		300.00
Institute of Statistics, Editorial Management		1000.00
ASA, ½ net profit Vols. 1–5 (5th and final payment)		432.20
Wm. Byrd Press		
Production of Biometrics, Vol. 10, 1954	$ 8582.00	
Offprints of Dec. 1953 to Sept. 1954	1232.80	9814.80
Total Expenses		$11681.73
Balance		$ 7992.34

Office of the Secretary-Treasurer

Income

Subscriptions, 1953—$369.50, 1954—$2913.75	$ 3283.25
Dues, 1953—$97.38, 1954—$1993.75	2091.13
Sustaining memberships, 1954	500.00
Back dues and subscriptions	41.50
Reprints, directories, back issues	14.45
Sale of excess equipment, bank charges	100.19
Overpayments ($122.16 less 55.91 for credits taken)	66.25
Total Income	$ 6096.77

Expenses

BIOMETRICS, 1954—$2262.50, 1953 and earlier—$49.82	$ 2312.32
Regional allotments	340.75
Salaries and special services	1438.52
Printing, stationary, supplies, etc.	159.56
Postage	136.48
Telephone	86.82
Depreciation	230.89
Total Expenses	$ 4705.34
Excess of Income over Expense	$ 1391.43

Balance Sheet as of December 31, 1954

Assets

Cash on hand (including petty cash)		$ 2908.73

Liabilities

Dues and Subscriptions for 1955	$ 218.50	
Surplus, Jan. 1, 1954 (excluding funds in transit)	1298.80	
Gain for period*	1391.43	$ 2908.73

*Owed against this gain but not paid in 1954—BIOMETRICS $1822.07, regional allotments $2.50; owing to this office for 1954 from Regional Treasurers and National Secretaries but not received in 1954—subscriptions $947.00, dues $261.75.

BIOMETRICS

VOLUME 11: 1955

Numbers 1–4

Published Quarterly by
THE BIOMETRIC SOCIETY
NEW HAVEN, CONN.
1955

STATEMENT REQUIRED BY THE ACT OF AUGUST 24, 1912, AS AMENDED BY THE ACTS OF MARCH 3, 1933, AND JULY 2, 1946 (Title 39, United States Code, Section 233) SHOWING THE OWNERSHIP, MANAGEMENT, AND CIRCULATION OF

BIOMETRICS published Quarterly at New Haven, Connecticut for January-December, 1955

1. The names and addresses of the publisher, editor, managing editor, and business managers are:

Name	*Address*
Publisher, The Biometric Society	52 Hillhouse Ave., New Haven, Conn.
Editor, Prof. Gertrude Cox	Box 5457, College Station, Raleigh, North Carolina
Managing Editor, Mrs. Sarah P. Carroll	Box 5457, College Station, Raleigh, North Carolina
Business Manager, None.	

2. The owner is: (if owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual member, must be given.)

Name	*Address*
The Biometric Society	52 Hillhouse Ave., New Haven, Conn.

3. The known bondholders, mortgagees, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages, or other securities are: (If there are none, so state.)

Name	*Address*
None	

4. Paragraphs 2 and 3 include, in cases where the stockholder or security holder appears upon the books of the company as trustee or in any other fiduciary relation, the name of the person or corporation for whom such trustee is acting; also the statements in the two paragraphs show the affiant's full knowledge and belief as to the circumstances and conditions under which stockholders and security holders who do not appear upon the books of the company as trustees, hold stock and securities in a capacity other than that of a bona fide owner.

5. The average number of copies of each issue of this publication sold or distributed, through the mails or otherwise, to paid subscribers during the 12 months preceding the date shown above was: (This information is required from daily, weekly, semiweekly, and triweekly newspapers only.).....

Sarah P. Carroll, Managing Editor

Sworn to and subscribed before me this 27 day of September 1955.

[SEAL]

Katherine C. McKimmon
(My commission expires April 18, 1957)

BIOMETRICS

Contents of Volume XI

ABSTRACTS 114, 240, 508, 535
NOTES 118, 519
QUERIES 111, 237, 395, 504
THE BIOMETRIC SOCIETY 115, 252, 556

ARMITAGE, P., A. W. DOWNIE and K. McCARTHY, *Abstract* #319 248
ARMITAGE, P., Tests for Linear Trends in Proportions and Frequencies 375
BAILY, J. L., Jr., *Abstract* #327 510
BAITSCH, H., *Abstract* #318 247
BARRACLOUGH, C. G., Statistical Analysis of Multiple Slope Ratio Assays 186
BAUER, R. K., *Abstract* #309 242
BINET, F. E. (*See* RICHARDSON, Mrs. G. L.)
BLACKITH, R. E., *Abstract* #323 250
BLISS, C. I., An Introductory Course in Biometry for Graduate Students In Biology 213
BLISS, C. I., *Abstract* #356 547
BOX, G. E. P., and P. V. YOULE, The Exploration and Exploitation of Response Surfaces: An Example of the Link Between the Fitted Surface and the Basic Mechanism of the System 287
BRANDT, A. E. and G. H. FLETCHER, *Abstract* #361 552
BRIEGER, F. G., *Abstract* #345 536
BRIER, O. G., M. SIQUEIRA and P. M. FREIRE, *Abstract* #358 549
CANSADO, E., *Abstract* #355 546
CHAPMAN, D. G. and RONALD PYKE, *Abstract* #326 508
CHARBONNIER, A., B. CYFFERS, D. SCHWARTZ and A. VESSEREAU, *Abstract* #362 553
CLARINGBOLD, P. J., Use of the Simplex Design in the Study of Joint Action of Related Hormones 174
CLARINGBOLD, P. J., *Abstract* #341 518
COCHRAN, W. G., Brief of Presidential Address 528
CONAGIN, ARMANDO and C. G. FRAGA, *Abstract* #347 539
COX, D. R., *Abstract* #320 248
COX, GERTRUDE M., *Abstract* #349 541
CYFFERS, B., (*See* CHARBONNIER, A.)
DEMPSTER, E. R., *Abstract* #344 535
DITCHBURNE, NELL, A Method of Analysis for a Double Classification Arranged in a Triangular Table 453
DOUGLAS, J. B., Fitting the Neyman Type A (Two Parameter) Contagious Distribution 149
DOWNIE, A. W. (*See* ARMITAGE, P.)
DRUCKREY, H., *Abstract* #310 243
DUNCAN, DAVID B., Multiple Range and Multiple *F* Tests 1
FIELDS, PAUL E., *Abstract* #325 508
FINKNER, A. L. (*See* HORVITZ, D. G.)
FINNEY, D. J. and G. C. VARLEY, An example of the Truncated Poisson Distribution 387
FINNEY, D. J., *Abstract* #357 548
FISHER, R. A., *Abstract* #343 535
FLEISHER, J. (*See* HORVITZ, D. G.)
FLETCHER, G. H. (*See* BRANDT, A. E.)
FRAGA, C. G. (*See* CONAGIN, ARMANDO)
FREIRE, PAULO MELLO (*See* BRIER, O. G.)

FRICK, JAMES W., *Abstract* #334 514
GAUL, H. and H. MUENZNER, *Abstract* #311 244
GENGERELLI, J. A., *Abstract* #332 513
GOEN, R. L. (*See* TATE, R. F.)
GOMES, F. PIMENTEL, *Abstract* #348 540
GRANT, DAVID A., *Abstract* #330 511
GREENHOUSE, SAMUEL W., *Abstract* #304 240
GRIDGEMAN, N. T. (*See* HOPKINS, J. W.)
GRIDGEMAN, N. T., The Bradley-Terry Probability Model and Preference Tasting 335
HAMAKER, H. C., Experimental Design in Industry 257
HANSEN, M. H. and J. STEINBERG, *Abstract* #353 545
HARMAN, HARRY H., *Abstract* #333 513
HAYMAN, B. I. and K. MATHER, The Description of Genic Interaction in Continuous Variation 69
HOLT, L. B., An Attempt to Derive a Mathematical Characterization of the Antigenicity of Diphtheria Prophylactic 83
HOPKINS, J. W. and N. T. GRIDGEMAN, Comparative Sensitivity of Pair and Triad Flavor Intensity Difference Tests 63
HORVITZ, D. G., J. FLEISHER and A. L. FINKNER, *Abstract* #306 241
HUDDLESTON, HAROLD F., An Inverted Matrix Approach for Determining Crop-Weather Regression Equations 231
HUET, A., D. SCHWARTZ, and A. VESSEREAU, *Abstract* #302 114
HUNTER, J. S. (*See* YOUDEN, W. J.)
IPSEN, JOHANNES, *Abstract* #305 240
IPSEN, JOHANNES, Appropriate Scores in Bio-Assays Using Death-Times and Survivor Symptoms 465
IRWIN, J. O., *Abstract* #359 550
JESSEN, RAYMOND J., Determining the Fruit Count on a Tree by Randomized Branch Sampling 99
KEITER, F., *Abstract* #312 245
KENADLL, M. G., Further Contributions to the Theory of Paired Comparisons 43
KIMMEL, HERBERT D., *Abstract* #329 511
KNEIP, P., *Abstract* #313 246
KOLIER, S., *Abstract* #314 246
KRAMER, CLYDE YOUNG, On the Analysis of Variance of a Two-Way Classification with Unequal Sub-Class Numbers 441
LESLIE, R. T., *Abstract* #338 516
LIDER, L. A., *Abstract* #335 514
LINDER, ARTHUR, *Abstract* #363 555
LOWRY, DOROTHY C., Variance Components with Reference to Genetic Population Parameters 136
LUDWIG, W., *Abstract* #315 246
LUDWIG, W., *Abstract* #316 247
MANCEAU, J. N., *Abstract* #360 552
MATHER, K. (*See* HAYMAN, B. I.)
McINTYRE, G. A., Design and Analysis of Two Phase Experiments 324
MERRIFIELD, PHILIP R., *Abstract* #331 512
MOTT, G. O., *Abstract* #352 544
MUENZNER, H. (*See* Gaul, H.)
NORTON, H. W., A Further Note on Missing Data 110
OLLAGNIER, M., *Abstract* #303 114
OUTHWAITE, ANNE D. and A. RUTHERFORD, Covariance Analysis as an Alternative to Stratification in the Control of Gradients 431
PANSE, V. G. (*See* SUKHATME, P. V.)
PEARCE, S. C., *Abstract* #346 538
PYKE, RONALD (*See* CHAPMAN, D. J.)
REEVE, E. C. R., The Variance of the Genetic Correlation Coefficient 357
RICHARDSON, Mrs. G. L. and F. E. BINET, *Abstract* #342 518

ROBERTSON, ALAN, Prediction Equations in Quantitative Genetics 95
DA ROCHA, G. L., *Abstract* #351 542
RUBIN, HERMAN, *Abstract* #337 516
RUTHERFORD, A. (*See* OUTHWAITE, ANNE D.)
SAMPFORD, M. R., *Abstract* #322 249
SCHULTZ, E. F., Jr., Rules of Thumb for Determining Expectations of Mean Squares in Analysis of Variance 123
SCHWARTZ, D. (*See* HUET, A.)
SCHWARTZ, D. (*See* CHARBONNIER, A.)
SIQUEIRA, M. (*See* BRIER, O. G.)
SMITH, CEDRIC A. B., *Abstract* #324 250
SPICER, C. C., The Theory of Bacterial Constant Growth Apparatus 225
SPRENT, P., A Note on Design and Analysis of Soil Insecticide Experiments 427
STEEL, R. G. D., An Analysis of Perennial Crop Data 201
STEINBERG, J. (*See* HANSEN, M. H.)
STORMONT, C. N., *Abstract* #336 515
SUKHATME, P. V., and V. G. PANSE, *Abstract* #354 545
TATE, R. F. and R. L. GOEN, *Abstract* #328 511
TEISSIER, GEORGES, Sur la Determination de l'axe d'un Nuage Rectilligne de Points 344
THOMPSON, W. A., Jr., The Relative Size of the Inter- and Intra-Block Error in an Incomplete Block Design 406
TWEEDIE, M. C. K., *Abstract* #307 242
VARLEY, G. C. (*See* FINNEY, D. J.)
VESSEREAU, A. (*See* HUET, A.)
VESSEREAU, A. (*See* CHARBONNIER, A.)
WALTER, E., *Abstract* #317 247
WATSON, G. S., *Abstract* #308 242
WATSON, G. S., *Abstract* #340 517
WILLIAMS, E. J., *Abstract* #339 517
YATES, F., *Abstract* #321 249
YOUDEN, W. J. and J. S. HUNTER, Partially Replicated Latin Squares 399
YOUDEN, W. J., *Abstract* #350 541
YOULE, P. V. (*See* BOX, G. E. P.)

INDEX, BIOMETRICS VOLUME 11

Acceptance area, see tests
Additivity, see analysis of variance, hypothesis, model, transformations test of, 110, 111, 400
Agronomy, 201, 231, 241, 422, 431, and see horticulture, long-term experiments, soil
Aliases, 483, 494
Allometry, 344, and see growth
All-or-none response, see quantal response
Allowances, Tukey's, 29
Animal breeding, 95, and see inbreeding
Animal experiments, litter mates in, 249, and see design of experiments
Anthropology, 245, and see medicine
Association, see contingency
Bacteriology, 225, 375
Balancing, see paired comparisons, randomized blocks, selection
Bartlett's test, see homogeneity of variance, multivariate tests
Bayes' distribution, 17
Between-block error, see interblock error
Binomial distribution, 388, 492, and see distributions, quantal response negative, 250, 394
Bioassay, 83, 240, 465, 481, and see biometry, dose response curve, pharmacology, quantal response, time response curve, tolerance, toxicology, transformations
  design, 190
  interpretation of, 246
  slope ratio, 186
  species transferability, 88
Biochemistry, 192
Biological assay, see bioassay
Biomathematics, see biometry
Biometrics, see biometry
Biometric Society, 115, 252, 556
Biometry, 84, 183, 213, 225, 243, 245, 246, 247, and see allometry, bioassay, growth, hypotheses, models, path coefficients, Weber-Fechner law
Biophysics, see biometry
Blood groups, 250
Calculation, see computation
Canonical analysis, 289
Chain blocks, see linked blocks
Characteristic function, 242, and see moments
Chemistry, see biochemistry, industrial research
Chi square, 338, 377, 397, 411, and see analysis of variance, contingency, fourfold table, goodness of fit, minimum chi square estimate
  analysis of, 393
Choice, see selection
Chromosomes, 244
Clinical experiments, 469
Combination of data, 95, 240, 249
Combination of tests, see tests
Comparisons, see analysis of variance, degrees of freedom, orthogonal polynomials, tests
  individual, 16
  orthogonal, 179, 454
  paired, 43, and see half-leaf method
Complementary genes, 73
Components, see correlation, covariance, variance
Computation, 151, 231, 481, 495, and see estimation, graphics, iteration, matrices, method of moments, method of steepest descent, Newton's method, recurrence formulas
  punched card, 114
Confidence limits, 189, and see tests
  for correlation, 362
  for difference of means, 505
  for mean, 504
  for relative potency, 485
  for variance components, 146
Confounding, 134, 399, and see analysis of variance, design of experiments, factorial experiments
Contagious distribution, Neyman's, 149
  truncated, 153
Contingency tables, 249, and see chi square, correlation, fourfold table

missing data, 242
trend in, 375
Contrasts, see comparisons
Correlation, see contingency, covariance, regression
between relatives, see statistical genetics
biological, 247
biserial, 384
canonical, 206
matrix, 346
multiple, 206
partial, 248
phenotypic, 97
rank, 380
Courses in statistics, see statistics curricula and syllabi
Covariance, see correlation, matrices, regression, variance
adjustment, 428
analysis of, 246, 429, 431
components of, 247
matrix, 182, 188, 207, 345, 411, 484
multiple, 439
Crossing over, 244, and see linkage
Curve fitting, see estimation, orthogonal polynomials, regression
Decision, 7, and see tests
regions, 24
Degrees of freedom, 16, 296, 395, and see comparisons
effective, 144
Design of experiments, 248, 249, 257, 427, 494, and see aliases, components of variance, confounding, factorial experiments, fractional replication, half-leaf method, incomplete blocks, interblock error, intrablock error, lattices, linked blocks, long-term experiments, nonfactorial, organolepsis, paired comparisons, parallelogram design, perennials, precision, randomization, randomized blocks, replication, sampling, tree crops, triangle test, Youden square
for locating optimum, 287
genetic, 140
judging, 57
nonfactorial, 183
simplex, 174
two-phase, 324
Differential equations, 225, 321
Discriminant function, 209, 240, 242, and see multivariate analysis, scores
Disproportionate subclasses, 441, and see analysis of variance, fitting constants, method of weighted squares of means
Distribution, see Bayes, binomial, characteristic function, chi square, contagious, goodness of fit, logarithmic normal, moments, normal, Poisson
beta, 204
exact, 384
fitting, see estimation
infectious, see contagious
Neyman's, see contagious
of chi square, 411, 479
of F, 411
tolerance, 174
truncated, 465
Dominance, 137
Dose response curve, 83, 176, 188, 243, 465, and see time response curve
Duo-trio test, 63
Duplicate genes, 73
Ecology, 387, 504, and see entomology, fish, genetics, populations
Efficiency, 149, 325, 388, 438, 478, and see estimation, maximum likelihood, sufficient statistics
sampling, 108, 242
Election, see judging
Endocrinology, 174
Entomology, 156, 250, 387, 427, and see toxicology
Equiprobable ellipse, 345
Epistasis, 71, 137
dominant, 73
recessive, 73
Errata, 113, 251, 334
Error,
gross, 278
mean square, 276, and see variance
regression, see analysis of covariance
residuals, 261, 278, 428
term, proper, 135, 334
theory of, see analysis of variance, least squares, maximum likelihood, models, rejection of data
Estimation, 149, 456, and see combina-

tion of estimates, efficiency, fitting constants, information, least squares, maximum likelihood, method of moments, minimum chi square estimate, populations, scores, Sheppard's correction, sufficient statistics, theory of error
inefficient, 391
of components of covariance, 123, 136
of potency, 83, 188, 485
of proportion, 250
of variance, 390
simultaneous, 149
Expectation, mathematical, 410, and see mean, moments
of mean square, 123
Experimental design, see design of experiments
Experimental units, 238, and see selection
Factor analysis, 346, and see multivariate analysis
Factorial experiments, 15, and see fractional replication, nonfactorial
analysis of, 114, 134
confounded, 399
Faulty analysis, 428
Fiducial limits, see confidence limits
Field experiments, see agronomy, design of experiments, long-term experiments, soil
Fish, 504, and see ecology, population management
Fitting constants, 443, 454, and see disproportionate subclasses, least squares, missing values
Fitting distribution, see estimation
Fitting regression line, see covariance, least squares, regression
Fourfold table, 249, and see chi square, contingency
Frequencies, trends in, 395
F test, 444, and see analysis of variance, range test, tests
approximate, 144, 395
errors heterogeneous, 395
interpretation of, see tests
multiple, 1, 32, 505
power of, 398, 444
significant, procedure when, 1
variances unequal, 395
Generating function, see characteristic function, moments
Genetic correlation, 139, 247, 357
Genetic model, see models
Genetics, 69, 250, and see animal breeding, blood groups, chromosomes, complementary genes, crossing over, dominance, duplicate genes, epistasis, heritability, inbreeding, linkage, overdominance, path coefficients
assortative mating, 140, 363
human, 242, 247
non-additive effects, 139
prediction, 95
statistical, 69, 95, 136, 245, 357
Genetic selection, 139
Genetic variance, 95, 137
Goodness of fit, 296, 389, 487, 500, and see chi square
Graeco-Latin square, 327
Graphics, 277, 466, and see computation
Grouping, 237
Growth, 225, and see allometry, biometry, logistic curve
Half-leaf method, 327
Hematology, 246, and see blood groups, serology
Heredity, see genetics
Heritability, 95, 144, 360
variance of estimated, 146
Horticulture, 99, 125, and see agronomy, perennials
Hypothesis, see models, tests, null, 16, 338, 395, 415
Immunology, 469, and see serology, tuberculin
Inbreeding, 139
Incomplete blocks, 336, 406, and see lattices, randomized blocks, Youden square
balanced, 56
partially balanced, 61, 419
Incomplete experiments, see missing values
Industrial research, 257, 406, 402, and see physical sciences
Inference, 123, and see tests
Information, 333, 404, 436, 469, 494,

and see analysis of variance, design of experiments, estimation, maximum likelihood, precision
interblock, 325
matrix, 483, 495
mean, 240
Interaction, 127, 202, and see analysis of variance, models
as error term, 114, 134
assumed zero, 15, 403, 454
genetic, 71
Iteration, 151, 251, 284, 319, 346, 483, and see computation, least squares, maximum likelihood
Judging, 43, and see organolepsis
coefficient of agreement, 48
Latin square, 325, 399, 438
in time, 111
missing values, 110
partially replicated, 399
Lattices,
balanced, 325
rectangular, 427
Least significant difference, 2
test, 34
Least squares, 143, 345, 358, 454, and see adjusted means, analysis of variance, estimation, fitting constants, matrices, normal equations, theory of error
Likelihood, 66, and see maximum likelihood
Likelihood ratio, 240, 338
Limnology, see fish
Linear dependence, 483, and see matrices
Linkage, 70, 138, 358, and see crossing over
Linked blocks, 417
partially, 419
Logarithmic normal distribution, 84, and see normal distribution
Logistic curve, 228, 337, and see growth, logit transformation, populations
Long-term experiments, 201, and see tree crops
Main effect, see analysis of variance, interaction
Matching, see balancing, paired comparisons
Mathematical biology, see biometry
Matrices, 51, 143, 175, 411, 468, 481, and see covariance matrix, information matrix, linear dependence, multivariate analysis
design, 176, 481
direct product of, 179
inverse, 182, 188, 231, 435, 495
tables of, 195
of coordinate functions, 483
powers of, 51
Maximum likelihood, 66, 242, 249, 251, 337, 390, 482, 496, and see efficiency, estimation, information, theory of error
iterative computation, 149, and see iteration
Mean, see moments
adjusted, 401, 427, 435, and see least squares
geometric, 84
weighted, 114, 325, 368
Medicine, 83, 375, and see clinical, genetics, hematology, immunology, pharmacology, physiology, Schick conversion rate, serology, toxicology, virology
Metameter, 465, and see scales
Metric, 312, and see scales
Minimum chi square estimate, 388
Missing values, 110, 249, and see fitting constants, mixed-up values, rejection of data
contingency tables, 242
impossible, 110
variance of, 110
Mixed-up values, 242, and see missing values
Models, see biometry, components of variance, hypothesis, regression, theory of error, transformations
mathematical, 86, 110, 174, 266, 407, 445
mixed, 123, 136, 407
multivariate, 202
probability, 335
statistical, 143, 390
Moments, 360
generating function, 360
method of, 149
Multivariate analysis, 201, 344, and see analysis of variance, discriminant

function, factor analysis, matrices
Newman-Keuls test, see multiple range test
Newton's method, 149
Nonfactorial experiments, 183
Normal distribution, see logarithmic normal distribution, normality
multivariate, 344
Normal equations, 231, 443, and see least squares, matrices
Normality, 278, and see normal distribution
departure from, 336
Organolepsis, 335, 406, and see judging, scores, triangle test
flavor, 63, 335
homogeneity of data, 65
Orthogonal functions, 411
Orthogonality, 190, 439, 484, and see comparisons
non-orthogonality, 441
Orthogonal polynomials, 271, 431
Overdominance, 141
Pairing, see balancing, paired comparisons
Palatability, see organolepsis
Parallelogram design, 483
Path coefficients, 98, 372
Percentages, see proportions
Perennials, 201, and see tree crops
Pharmacology, 243, and see toxicology
Physical science, 238, 290, and see industrial research
Physiology, 86, and see endocrinology, medicine, threshold
Poisson distribution, 151, 248, 495
moment ratio tables, 157
truncated, 387
Pooling, 249, 362, 410, 472
Populations, see distributions, fish, statistical genetics
management, 225, and see ecology
Precision, 429, and see information
Prediction, see regression
genetic, 95
Preferences, see organolepsis
Proportions, 250, and see binomial
trend in, 375
Protection level, 13
consistent, 14
Quantal response, 481, and see bioassay, models
Quantification, see scales
Randomization, 241, 325, 493, and see selection
Randomized blocks, 14, 110, 111, 431, and see balancing, incomplete blocks, interblock error
Range,
shortest significant, 5
significant studentized, 5
table of, 3
test, multiple, 1, 26, 212
interpretation of, 6
power of, 2
Tukey's, 31
Ranks, 43, 335, 407, and see scores, transformations
of means, 7
tied, 47
Recurrence formulas, 150
Regression, see adjusted means, canonical analysis, correlation, covariance, orthogonal polynomials, trend
adjustment, 428
analysis, 180, 264, 376, 482, and see analysis of covariance
coefficient, 85, 270, 435, 484
biased, 483
crop-weather, 231
equation, 231
homogeneity of, 246
independent variable affected by treatments, 430
interpretation of, 379
line, 344
multiple, 188, 231, 344
weighted, 368
Rejection of data, 278, and see gross errors, missing values, selection
Replication, 269, 324, 439
fractional, 15, 399, and see aliases
Residuals, see errors
Response surface, 183, 287
Result-guided procedures, 239
Sample size needed, 36
Sampling, 125, and see components of variance, design of experiments
efficiency, 108
error, see variance

nonrandom, 241
random, 99, 241
stratified, 108, 431
Scales, 76, and see metric, scores
Scheffé's test, 35
Schick conversion rate, 85
Scores, 240, 335, 376, 465, and see discriminant functions, organolepsis, ranks, scales
iterated, 50
tournament, 49
Selection, see balancing, design of experiments, genetic selection, randomization, rejection of data, sampling
of characters, 243
of experimental units, 89, 238
of method of statistical analysis, 239
of scores, 378
of significance level, 14, 277
Sensory tests, see organolepsis
Serology, 83, 114, and see blood groups, hematology, immunology
Sheppard's correction, 237
Significance, 11, and see selection regions, 24
Simultaneous equations, see matrices
Sociology, 250
Soil science, 427
Squariance, 338, and see variance
Standard deviation, see variance
Standard error, see variance
Statistical control, see analysis of covariance
of gradient, 431
Statistics curricula and syllabi, 213, 254
Steepest descent, method of, 288, 318
Stochastic process, 247
Structural analysis, 126
Student's t, see t test
Subjective evaluation, see judging, organolepsis
Successive approximation, see iteration
Sufficient statistics, 149, and see efficiency, estimation
Survival curve, see time response curve
Survival time, see time response curve
Tables, miscellaneous, 37, 38, 72, 159, 177, 195
Taste tests, see organolepsis
Teaching of statistics, 118, 213, and see statistics curricula and syllabi
Tests, see analysis of variance, chi square, confidence limits, F test, goodness of fit, least significant difference, likelihood ratio, multiple range test, null hypothesis, protection level, ranks, rejection of data, result-guided, Scheffé's, significance, triangle test, t test
choice of, 383, 435
combination of, 201, 249
efficiency, 249
independent, 15
interpretation of, 6, 211, 343
multiple, 1
multiple comparisons, 33
multivariate, 204
normal deviate, multiple, 14
normality, 278
of significance,
Type I error, 11
Type II error, 14
of significance of
correlation, 362
extreme deviate, 40
interaction, 278
largest difference, see multiple range test
regression, 375
optimum, 17
power of, 9, 64
range, 1
rank correlation, 380
results suggested by data, see result-guided
Theory, see biometry, hypothesis, model
Threshold, 335, 481
Time response curve, 249, 465, and see dose response curve
Tolerance, 174
Tournaments, scoring, 49
Toxicology, 174, 243, and see bioassay, pharmacology
Transformations, 249, 465, 481, and see additivity, analysis of variance, bioassay, canonical analysis, logistic curve, matrices, models
angular, 483, 498
linear, 175, 207
logarithmic, 248

log dose, table of, 177
logit, 337, 481
normit, 337
orthogonal, 175
probit, 84, 174, 337, 481
squared hyperbolic secant, 337
square root, 327, 495
Tree crops, 99, 125, and see long-term experiments
Trend, 375, and see regression
Triangle test, 63
T test, see confidence limits
grouped data, 237
multiple, 24
power of, 248
symmetric three-decision, 19
Tuberculin, 114
Unequal subclasses, 441
Uniformity data, 430
Variance, see covariance, *F* test, genetics, mean square error, Sheppard's correction, squariance
analysis of, 123, 178, 190, 201, 259, 400, 427, 441, 499, and see additivity, chi square, combination of data, components of variance, degrees of freedom, disproportionate subclasses, error, fitting constants, *F* test, interblock error, least significant difference, least squares, long-term experiments, missing values, models, multiple *F* test, multivariate analysis, orthogonal polynomials, pooling, regression, structural analysis, tests, transformations, uniformity data
computation of, 111, 114, 179
fixed effects, 123
interpretation of, 139, 259, 276, 290, 410, 428
mixed model, 123, 136
of triangular array, 453
random effects, 123, 136
approximate, 198, 359
asymptotic, 391
components, 123, 136, 204, 237, 329, 416, and see components of covariance
computation of, 143
confidence limits for, 144
variance of, 144
expectation, 123, 141
homogeneity of, 232, 340, 395
interblock, 406
intrablock, 406
matrix, see covariance matrix
negative, 441
of adjusted mean, 436
of difference of adjusted means, 402
of estimate, 251
of estimate of heritability, 146
of genetic correlation, 357
of regression coefficient, 231, 367, 378, 435, 484
phenotypic, 137, 360
ratio, confidence limits for, 146, 407
sampling, 105
theoretical, 123
Variate, canonical, 206
Virology, 248, 326, and see half-leaf method
Weber-Fechner law, 335
Weighted squares of means, method of, 443
Weighting, 102, 142, 396, 444, 489, 499
Youden squares, 57
Zoology, 352
Z test, see *F* test